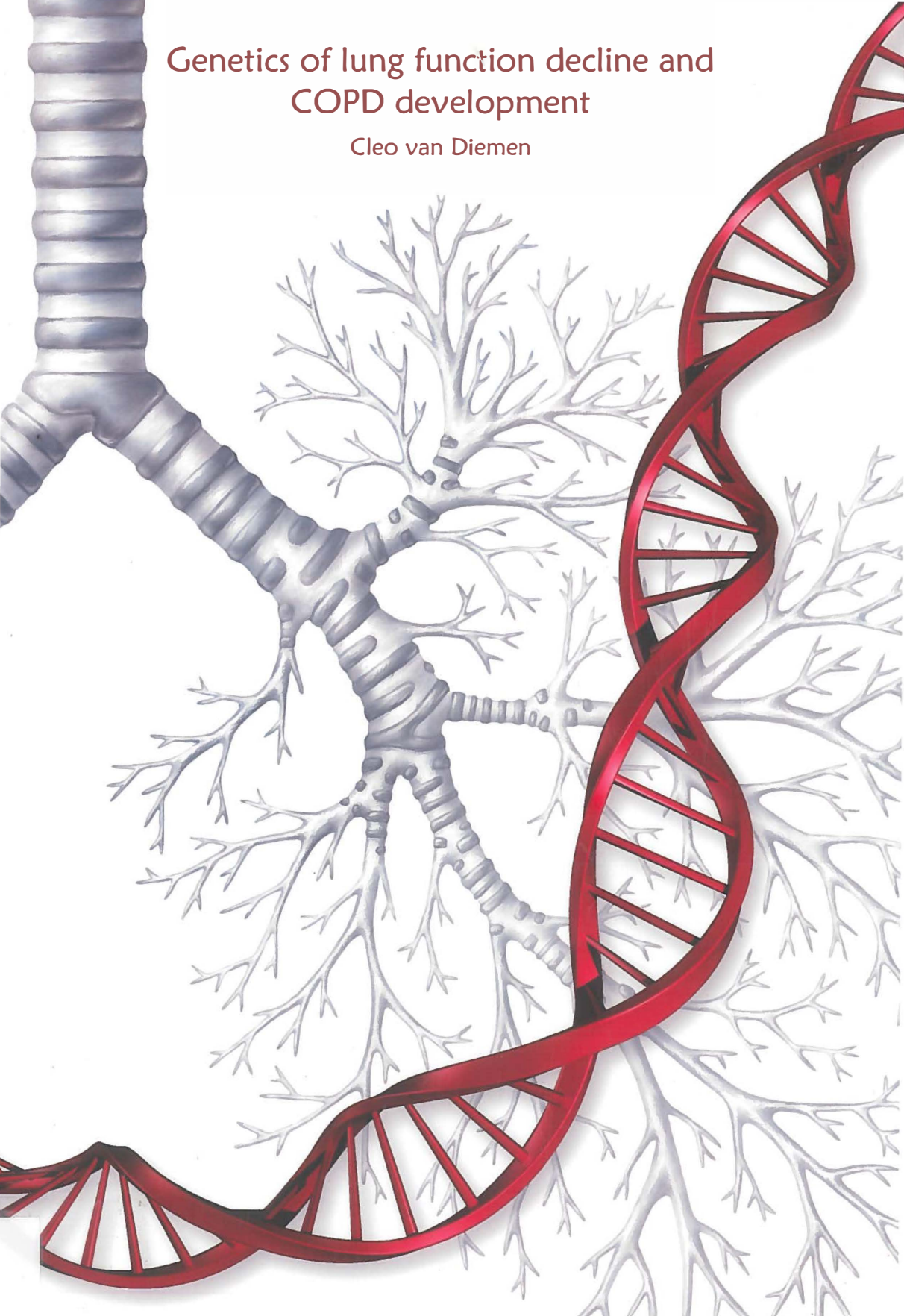
Genetics of lung function decline and COPD development

Diemen, Cleo Catharina van

Genetics of lung function decline and COPD development

Cleo van Diemen



Genetics of lung function decline and COPD development

Cleo van Diemen

C.C. van Diemen
Genetics of lung function decline and COPD development

Thesis University of Groningen with summary in Dutch

The studies described in this thesis were financially supported by the Dutch Asthma Foundation and the University of Groningen

Printing of this thesis was financially supported by: Groningen University Institute for Drug Exploration (GUIDE), Stichting Astmabestrijding, Rijksuniversiteit Groningen, Boehringer Ingelheim BV, Meda Pharma BV, Novartis Pharma BV, AstraZeneca BV.

Cover design by Jeroen Advocaat (www.jeroenadvocaat.com)

Printed by Gildeprint, Enschede

ISBN: 978-90-367-3329-8

LMB.

RIJKSUNIVERSITEIT GRONINGEN

GENETICS OF LUNG FUNCTION DECLINE AND COPD DEVELOPMENT

Proefschrift

ter verkrijging van het doctoraat in de
Medische Wetenschappen
aan de Rijksuniversiteit Groningen
op gezag van de
Rector Magnificus, dr. F. Zwarts,
in het openbaar te verdedigen op
woensdag 23 april 2008
om 16.15 uur

door

Cleo Catharina van Diemen
geboren op 13 november 1979
te Amersfoort

Promotor:	Prof. dr. Dirkje S Postma
Copromotor:	Dr. H Marike Boezen
Beoordelingscommissie:	Prof. dr. Gerard H Koëter Prof. dr. Peter D Paré Prof. dr. Cisca Wijmenga

Paranimfen: Hiltje Oude Luttikhuis
Mireille Edens

CONTENTS

LIST OF ABBREVIATIONS

- A : Adenine
- AAT : α1-Antitrypsin
- ADAM33 : A Disintegrin And Metalloprotease
- BAL : Bronchoalveolar Lavage
- BHR : Bronchial Hyperresponsiveness
- C : Cytosine
- Cm : Centimorgan
- COPD : Chronic Obstructive Pulmonary Disease
- CI : Confidence Interval
- DF : Degrees of Freedom
- DNA : Deoxyribonucleic Acid
- ECM : Extracellular Matrix
- EGF : Endothelial Growth Factor
- FEF : Forced Expiratory Flow
- FEV_1 : Forced Expiratory Volume in one second
- FEV_1 % pred : FEV_1 as percentage of predicted value
- FVC : Forced Vital Capacity
- G: Guanine
- GLUCOLD : Groningen Leiden Universities Corticosteroids in Obstructive Lung Disease
- GOLD : Global initiative for chronic Obstructive Lung Diseases
- GRIP : Genetic Research in Isolated Population
- GST : Glutathione S-Transferase
- IL : Interleukin
- LPC : Lysophosphatidylcholine
- LD : Linkage Disequilibrium
- LME : Linear Mixed Effects
- LOD : Logarithm of the Odds
- LP : Lysophospholipids
- LTX : Lung Transplantation
- LPC : Lysophosphatidylcholine
- MMP : Matrix Metalloprotease
- OR : Odds Ratio
- PC : Phosphatidylcholine
- PC_{20} : Provocative Concentration
- RNA : Ribonucleic Acid
- SD : Standard Deviation
- SE : Standard Error

- SFTP : Surfactant Protein (gene annotation)
- SNP : Single Nucleotide Polymorphism
- SOD : Superoxide Dismutase
- SP : Surfactant Protein
- T : Thymine
- TGF-β1 : Transforming Growth Factor-β1
- TIMP : Tissue Inhibitor of Matrix Metalloprotease
- TNF-α : Tumour Necrosis Factor-α
- VC : Vital Capacity

CHAPTER 1

General Introduction

Chronic Obstructive Pulmonary Disease

Chronic Obstructive Pulmonary Disease (COPD) is a disease entity with two identifiable types of lung diseases, i.e. chronic bronchitis and emphysema. COPD is characterized by airway obstruction, airway inflammation and degradation of lung tissue. Chronic bronchitis is hallmarked by hyperplasia and hypertrophy of the goblet cells (mucous gland) of the airway, resulting in an increase in secretion of mucus which contributes to airway obstruction and respiratory symptoms like chronic cough and phlegm. Emphysema is defined histologically as the enlargement of the alveolar spaces at the end of the terminal bronchioles, with destruction of alveolar walls. The latter reduces the surface area available for gas transport which ultimately leads to insufficient oxygen uptake in the bloodstream and provision to body tissues, leading to shortness of breath (dyspnea) and exercise limitation in more severe cases of COPD. Most COPD patients display characteristics of both chronic bronchitis and emphysema, although the extent of both characteristic features may vary between patients.

Central to the definition of COPD is a reduced level of lung function and this provides the most important and robust phenotype of COPD. The Global Initiative for Chronic Obstructive Pulmonary Disease (GOLD) has classified the severity of COPD using the spirometric measures Forced Expiratory Volume in 1 second (FEV_1) and the Forced Vital Capacity (FVC) as presented in table 1 [1]. COPD is a progressive disease that is irreversible, and still no medication is available to prevent the accelerated decline in lung function. This is important since the decline in lung function predicts progression of the disease and overall mortality [2;3].

Table 1: Spirometric classification of COPD severity based on post-bronchodilator lung function measurements

GOLD stage	Characteristics
I: Mild COPD	• $FEV_1/FVC < 70\%$ • $FEV_1 \geq 80\%$ predicted
II: Moderate COPD	• $FEV_1/FVC < 70\%$ • $50\% \leq FEV_1 < 80\%$ predicted
III: Severe COPD	• $FEV_1/FVC < 70\%$ • $30\% \leq FEV_1 < 50\%$ predicted
IV: Very Severe COPD	• $FEV_1/FVC < 70\%$ • $FEV_1 < 30\%$ predicted or $FEV_1 < 50\%$ predicted plus chronic respiratory failure

Abbreviations: FEV_1: Forced Expiratory Volume in one second; FVC: Forced Vital Capacity; respiratory failure: arterial pressure of oxygen (PaO_2) less than 8.0 kPa (60 mm Hg) with or without arterial pressure of CO_2 ($PaCO_2$ greater than 6.7 kPa (50 mm Hg) while breathing air at sea level.

The main risk factor for COPD is smoking: 10-20% of smokers develop COPD, and virtually all COPD patients have smoked. Smoking induces chronic inflammation of the airways and lung tissue, leading to airway remodeling and degradation of lung tissue. It has been shown that this inflammatory process persists after smoking cessation [4;5]. Microscopically there is infiltration of the airway walls with inflammatory cells, particularly neutrophils

and T and B cells [6-9]. Chronic inflammation due to persistent smoking is followed by scarring and remodeling that thickens the airway walls resulting in narrowing of the small airways. Airway remodeling may be accompanied by metaplasia and fibrosis of the lower airways. It is generally accepted that several mechanisms contribute to the pathology of COPD such as a disturbed oxidant-antioxidant balance, a disturbed protease-antiprotease balance, cell apoptosis, and impaired tissue repair. These mechanisms interact and inflammation and tissue remodeling occur as a consequence of these phenomena.

Oxidants are reactive oxygen species that enter the lungs continuously by inhaling air. Tissue oxidative stress can normally be counteracted by exogenous antioxidant uptake from food and by endogenously present antioxidants like antioxidant enzymes that can detoxify reactive oxidant species. Cigarette smoke and air pollution contain excess oxidants that cannot be cleared immediately from the lung, leading to an imbalance between oxidants and antioxidants, so-called oxidative stress [10-12]. Elevated oxidative stress levels in induced sputum have been associated with the severity of COPD [13]. Oxidative stress amplifies inflammation by activating the transcription factor NF-κB, with subsequent increase of pro-inflammatory mediators [14]. When activated, the inflammatory cells produce oxidants themselves as well, thereby giving rise to a loop of oxidative stress and chronic inflammation of the airways.

Emphysematous lungs display breakdown of connective tissue components in the lung parenchyma by proteases such as neutrophil elastase, matrix metalloproteases and cathepsins [15-21]. Cigarette smoke induces release of proteases as well as their inhibitors, anti-proteases, to neutralize the protease activity and thus tissue breakdown. It has been proposed that in smokers who develop COPD, the balance between proteases and antiproteases is disturbed with subsequent parenchymal injury. The best example for this imbalance is α1-antitrypsin deficiency, which is caused by a polymorphism in the gene coding for α1-antitrypsin, resulting in a lack of functional α1-antitryspin protein in carriers of this polymorphism. Loss of the α1-antitrypsin protein results in less inhibition of the protease elastase, which may lead to alveolar destruction and development of emphysema even without cigarette smoke exposure.

Destruction of extracellular matrix (ECM) can not only be caused by excess protease activity, but may also be the consequence of insufficient repair. It has previously been shown that elastin and collagens and certain proteoglycans, i.e. decorin and biglycan, are less present in lung tissue from COPD patients compared to controls [22-24]. The Transforming Growth Factor-β1 (TGF-β1)-Smad pathway may play a central role in this process since it has been shown to induce ECM protein production. In COPD patients the TGFβ-Smad pathway is aberrantly expressed compared to healthy controls [25].

Given the fact that COPD is a heterogeneous disease, it is easy to understand that inflammatory patterns with different intensity of the various cell components have been described in the literature. Generally, COPD patients have a marked increase in activated neutrophils in induced sputum, BAL fluid and bronchial biopsies compared to smokers

without airway obstruction [6-8]. Increased numbers of $CD8^+$ lymphocytes have been shown throughout the airway walls and in lung tissue of COPD patients [26-28]. More recently, studies focus on the role and implication of B cells in COPD pathogenesis [9;29;30]. Eosinophils may play a role in COPD pathogenesis as well, since blood eosinophilia itself is an independent predictor of reduced lung function levels and accelerated lung function decline and indicates worse prognosis in chronic bronchitis [31-33].

There is an increase in apoptotic alveolar epithelial and endothelial cells in the lungs of COPD patients [34;35]. Since this is not counterbalanced by an increase in proliferation of these structural cells, the net result is disruption of epithelial integrity, destruction of lung tissue and the development of emphysema [34].

Genetics of Chronic Obstructive Pulmonary Disease

The level of lung function and the development of COPD are both environmentally and genetically determined. In the past decades, epidemiological studies have analyzed the environmental determinants extensively and there is now abundant evidence that the major environmental factor contributing to COPD development and excess lung function decline in the western world is cigarette smoking. It is, however, less clear why some smokers develop COPD and other smokers do not. This may be either due to additional environmental factors (e.g. less intake of oxidants in food) or to genetic factors. Clearly, it is of major importance to understand why some people are genetically susceptible to develop COPD while others are not, since such knowledge may provide new insights in the disease, potentially leading to better treatments. The genetic determinants are complicated to study, since COPD is a disease that is merely expressed at older ages. Therefore, the 'traditional' types of genetic studies, like those performed within families, are difficult to perform, because the parents of individuals with COPD often already have died and the children of individuals with COPD are likely to be too young to manifest airway obstruction already at that age. Case-control studies investigating COPD patients and healthy controls are generally used to study the genetics of COPD. So far these studies have been of small sample size and the definition of COPD has been variable. As a consequence, many studies cannot confirm previously reported genetic associations. Furthermore, the disease is only investigated when symptoms are present, without acknowledging the different patterns through which lung function loss may have evolved, i.e. an abnormally high rate of decline in lung function, a sub maximally attained level of lung function, or an abnormally early age of onset of decline. Genetic factors may affect any one, or a combination of these different patterns of lung function loss. Therefore, it is important to study multiple measurements of lung function in a well-characterized cohort that is followed up throughout life.

In our studies, we have chosen to study genetic effects on robust phenotypes, namely level of FEV_1 and rate of decline in FEV_1. A short introduction on genetics and the type of genetic variation (Single Nucleotide Polymorphisms (SNPs)) that we study is presented in the next paragraphs.

Brief introduction in genetics

Each human cell nucleus contains a set of chromosomes, making up the genome. The human genome consists of 23 pairs of chromosomes, consisting of 1 pair of sex chromosomes (X and Y). The DNA (deoxyribonucleic acid) molecule consists of two ribbon-like strands that wrap around each other, resembling a twisted ladder, often described as a double helix. The DNA is contained in tightly coiled packets called chromosomes, found in the nucleus of every cell. Chromosomes consist of the double helix of DNA wrapped around proteins, called histones (see figure 1).

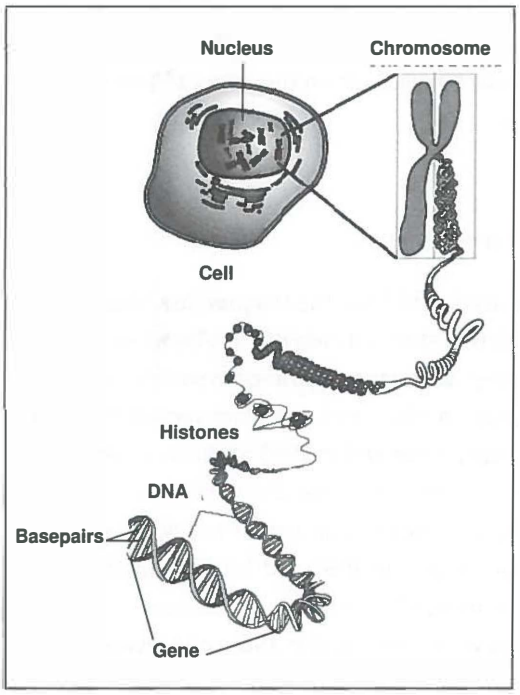


Figure 1: Schematic view of a cell with the nucleus containing the chromosomes. *The unwrapped chromosomes contain strands of DNA around histones.*

The double helix is made up of repeating units called nucleotides, who are composed of one sugar-phosphate molecule (the linear strands or outer rails of the ladder) and one base. DNA consists of two nucleotide strands joined by weak chemical bonds between the two bases, forming basepairs. A basepair is one step on the ladder of the DNA. The bases are called A (for adenine), C (for cytosine), T (for thymine) and G (for guanine). These bases always pair up in the following way: **A+T** and **C+G**.

A single strand of DNA is made of letters:
ATGCTCGAATAAATGATTTGA
The letters make words of three basepairs that each code for an amino acid:
ATG CTC GAA TAA ATG ATT TGA
The words make sentences, the genes:
<ATG CTC GAA TAA> <ATG ATT TGA>
Amino acids are the building blocks of proteins. Genes thus contain the information to produce proteins. Proteins are required for the structure, function, and regulation of the body's cells, tissues, and organs.

A mutation or polymorphism is a change in the DNA basepairs of a gene or an alteration in the chromosomes. Polymorphisms are common differences in the sequence of DNA, occurring in at least 1% of the population, whereas so-called 'mutations' occur in less than 1% of the population. Most DNA variation is neutral (not beneficial or harmful), but harmful sequence changes sometimes do occur. Changes within genes can result in proteins that do not work normally or do not work at all. In general, such changes can contribute to disease development, or diseases severity, and can also affect the response to a medicine. In the lungs, changes in genes that are involved in e.g. the oxidant/anti-oxidant balance and the protease/anti-protease balance within the lung can therefore be detrimental for lung tissue damage and repair.

THE VLAGTWEDDE/VLAARDINGEN COHORT STUDY

For most studies described in this thesis, we used data from the Vlagtwedde/Vlaardingen cohort study, an extensive longitudinal population study carried out in Vlagtwedde and Vlaardingen in the Netherlands. The first surveys consisted of a random sample of both areas of all men and women aged 40 to 64 years in 1965, and a random sample of all men and women aged 15 to 39 years in 1967 at Vlagtwedde and in 1969 at Vlaardingen. The cohort for the longitudinal study consisted of the youngest men and women who participated in 1965 (40 to 44 years old in the Vlagtwedde population and 40 to 54 years old in the Vlaardingen population), and all participants in 1967 and 1969 (aged 15 to 39 years). After the baseline surveys the cohort participated in follow-up surveys approximately every 3 years. The final surveys were organized in 1989 in Vlagtwedde and in 1990 in Vlaardingen. All surveys were carried out during the month of October. This large cohort of male and female smokers and non-smokers has thus been followed over a 25 yr period. This did not only allow us to study the contribution of genetic variations to decline in lung function but also genes involved in previously well established risk factors for lung function course and decline in this population, i.e. bronchial hyperresponsiveness (BHR) and blood eosinophilia. In 2003/2004, DNA was isolated from 2467 subjects of the Vlagtwedde/Vlaardingen cohort participating in the last survey in 1989/1990. In this thesis we have used data from 1390 subjects, selected on DNA availability (>1500 ng isolated DNA available).

Although the Vlagtwedde/Vlaardingen population is the main focus of research in this thesis, some of the studies were performed using a second population-based cohort, respectively 3 distinct COPD patient populations. The Doetinchem cohort is a random sample of the larger MORGEN cohort [36;37]. This population-based cohort has been followed up for 5 years, allowing us to study genetic effects on the decline in lung function and development of COPD in the general population.

The first COPD patient population consisted of individuals who were included in the GLUCOLD study (Groningen Leiden Universities Corticosteroids in Obstructive Lung Disease), a randomized control trial to analyze short versus long-term effects of steroids with and without bronchodilator in patients with mild to moderate COPD (GOLD stage II and III). A unique feature of this cohort is that bronchial biopsies were collected from all participants, as well as sputum, blood and lung function measures.

The second COPD patient population consisted of a cohort of patients with very severe COPD (GOLD stage ≥ III). We retrieved longitudinal pulmonary function measurements before the patients underwent lung transplantation (LTX) because of COPD. Approximately half the patients in the cohort have α1-antitrypsin deficiency, which enables us to compare genetic effects between patients with and without α1-antitrypsin deficiency.

The third COPD study population was derived from a genetically isolated population from the south of the Netherlands as part of the Genetic Research in Isolated Populations study (GRIP). Since the genetic variation in such an isolate is limited, it is thought that one can more easily identify genes involved in disease is such a population than in a general population. An overview and brief description of the different study populations is presented in table 2.

Table 2: Study populations described in this thesis

	Vlagtwedde/ Vlaardingen	Doetinchem	GLUCOLD	LTX	GRIP
Selection criteria	General population	General population	COPD patients GOLD stage ≥ II	COPD patients GOLD stage ≥ III who underwent lung transplantation	COPD patients GOLD stage ≥ I from a genetic isolate
Nr of subjects	1390	1152	114	79	106
Type of data	Longitudinal lung function data, BHR, blood eosinophils and symptoms for 25 years follow-up	Longitudinal lung function data and symptoms for 5 years follow-up	Cross-sectional measures of inflammation in sputum, blood and bronchial biopsies, BHR, lung function	Longitudinal lung function data prior to transplantation; patients with and without α1-antitrypsin deficiency	Cross-sectional lung function data in a genetically isolated population with genealogy data

Abbreviations: GLUCOLD Groningen Leiden Universities Corticosteroids in Obstructive Lung Disease; LTX lung transplantation; GRIP Genetic Research in Isolated Populations; COPD Chronic Obstructive Pulmonary Disease; GOLD Global initiative for Obstructive Lung Diseases; BHR Bronchial Hyperresponsiveness

AIMS OF THIS THESIS

In this thesis, studies are described that investigate the effects of SNPs in genes on lung function level and decline in lung function, COPD development and risk factors for COPD, i.e. BHR and blood eosinophilia. These genes are involved in oxidative stress, inflammation, detoxification of cigarette smoke and destruction of ECM components, like proteases. The specific subjects covered are described in the following paragraphs.

Chapter 2 gives an overview on the genetic epidemiology of reduced lung function, in which we present multiple different designs to study the role of genetics in development of COPD and reduced lung function.

Chapters 3, 4 and ***5*** present the studies on one gene, i.e. *A Disintegrin And Metalloprotease 33* (*ADAM33*) in 3 different populations. *ADAM33* has previously been identified as candidate gene for asthma in a whole genome wide screen, and has been associated with excess FEV_1 decline in asthmatics. In ***chapter 3***, the hypothesis is tested whether *ADAM33* may not only be associated with excess FEV_1 decline in asthmatics, but that this association might be observed in the general population as well. Since *ADAM33* has previously also been associated with BHR in asthmatics and the majority of COPD patients have BHR, we wondered whether *ADAM33* plays a role in BHR within a group of COPD patients. Therefore, in ***chapter 4***, the effect of SNPs on BHR severity and inflammation was studied in sputum and bronchial biopsies in 114 patients with mild to moderate COPD from the GLUCOLD study. Finally, in ***chapter 5*** we studied whether *ADAM33* SNPs also plays a role in excess FEV_1 decline within COPD patients. The effect of *ADAM33* SNPs on FEV_1 decline was therefore investigated in 79 patients with severe COPD with and without α1-antitrypsin deficiency with longitudinal lung function data prior to lung transplantation.

In ***chapter 6***, the association was studied between SNPs in *matrix metalloprotease* (*MMP*) genes and their tissue inhibitors and COPD and accelerated FEV_1 decline. As described previously, one of the most recognized theories behind development and progression of COPD is the disturbed balance of proteases and anti-proteases. This may be genetically determined.

In ***chapter 7***, we studied the role of *TGF-β1* and *decorin* in COPD development and lung function decline. TGF-β1 is a cytokine that is an important mediator in tissue proliferation and differentiation. From previous studies it is known that levels of TGF-β1 are increased in sputum and lung tissue from COPD patients, whereas levels of its inhibitor, decorin, are decreased. We analyzed whether this imbalance between TGF-β1 and decorin has a genetic origin and whether SNPs in their genes affect lung function decline and COPD development in the general population.

In ***chapter 8***, the focus lies on the effects of genetic variation in 3 antioxidant enzymes on lung function and how they interact with smoking and gender: Glutathione S-Transferases (GST) M1, P1 and T1. The effect of deletion alleles of *GSTM1* and *GSTT1* and 2 SNPs in *GSTP1* on the level and course of lung function was studied in the Vlagtwedde/Vlaardingen population.

Chapter 9 describes the effects of polymorphisms in the antioxidant enzymes superoxide dismutase 2 and 3 on level of lung function, lung function decline and bronchial hyperresponsiveness in the Vlagtwedde/Vlaardingen population.

In ***chapter 10***, we investigated the effects of SNPs in surfactant protein genes on decline in lung function and COPD development, and on well-known risk factors of COPD, i.e. BHR and blood eosinophilia. The pulmonary surfactant is important in alveolar structuring, bronchial clearance, and the innate immunity and additionally attracts eosinophils to the lungs. Therefore, changes in surfactant quantity or composition may contribute to lung function decline and development of COPD. Since a major part of the pulmonary surfactant consists of lipids, we analyzed surfactant lipids and their degradation products in sputum of COPD patients and controls and studied their association with lung function.

Chapter 11 describes a candidate gene study for level of airflow limitation in patients with mild COPD from a recently isolated population from the southwestern area of the Netherlands as part of the GRIP study. It has advantages to try and identify risk genes in populations that are homogeneous, like genetically isolated populations. Due to the small number of founders in genetically isolated populations, the total gene pool and therefore the number of different genes involved in a trait is limited. We analyzed the effects of 33 SNPs in 13 candidate genes for COPD on level of lung function and aimed to replicate the significant associations in COPD patients selected from the Vlagtwedde/Vlaardingen population.

Chapter 12 concludes with a summary and general discussion of the studies described in this thesis. Additionally, we describe future perspectives for further research on genetics of lung function and COPD development.

REFERENCES

1. From the Global Strategy for the Diagnosis, Management and Prevention of COPD, Global Initiative for Chronic Obstructive Lung Disease (GOLD) 2006. Available from: http://www.goldcopd.org.
2. Ekberg-Aronsson M, Pehrsson K, Nilsson JA, *et al*. Mortality in GOLD stages of COPD and its dependence on symptoms of chronic bronchitis. *Respir Res* 2005;6:98.
3. Sin DD, Man SF. Chronic obstructive pulmonary disease as a risk factor for cardiovascular morbidity and mortality. *Proc Am Thorac Soc* 2005;2(1):8-11.
4. Gamble E, Grootendorst DC, Hattotuwa K, *et al*. Airway mucosal inflammation in COPD is similar in smokers and ex-smokers: a pooled analysis. *Eur Respir J* 2007;30(3):467-71.
5. Willemse BW, ten Hacken NH, Rutgers B, *et al*. Effect of 1-year smoking cessation on airway inflammation in COPD and asymptomatic smokers. *Eur Respir J* 2005;26(5):835-45.
6. Baraldo S, Turato G, Badin C, *et al*. Neutrophilic infiltration within the airway smooth muscle in patients with COPD. *Thorax* 2004;59(4):308-12.
7. Keatings VM, Collins PD, Scott DM, *et al*. Differences in interleukin-8 and tumor necrosis factor-alpha in induced sputum from patients with chronic obstructive pulmonary disease or asthma. *Am J Respir Crit Care Med* 1996;153(2):530-4.
8. Lacoste JY, Bousquet J, Chanez P, *et al*. Eosinophilic and neutrophilic inflammation in asthma, chronic bronchitis, and chronic obstructive pulmonary disease. *J Allergy Clin Immunol* 1993;92(4):537-48.
9. Gosman MM, Willemse BW, Jansen DF, *et al*. Increased number of B-cells in bronchial biopsies in COPD. *Eur Respir J* 2006;27(1):60-4.
10. Rytila P, Rehn T, Ilumets H, *et al*. Increased oxidative stress in asymptomatic current chronic smokers and GOLD stage 0 COPD. *Respir Res* 2006;7:69.
11. Rahman I. Oxidative stress in pathogenesis of chronic obstructive pulmonary disease: cellular and molecular mechanisms. *Cell Biochem Biophys* 2005;43(1):167-88.
12. Rahman I, Morrison D, Donaldson K, *et al*. Systemic oxidative stress in asthma, COPD, and smokers. *Am J Respir Crit Care Med* 1996;154(4 Pt 1):1055-60.
13. Kanazawa H, Yoshikawa J. Elevated oxidative stress and reciprocal reduction of vascular endothelial growth factor levels with severity of COPD. *Chest* 2005;128(5):3191-7.
14. Rahman I, Marwick J, Kirkham P. Redox modulation of chromatin remodeling: impact on histone acetylation and deacetylation, NF-kappaB and pro-inflammatory gene expression. *Biochem Pharmacol* 2004;68(6):1255-67.
15. Cataldo D, Munaut C, Noel A, *et al*. MMP-2- and MMP-9-linked gelatinolytic activity in the sputum from patients with asthma and chronic obstructive pulmonary disease. *Int Arch Allergy Immunol* 2000;123(3):259-67.
16. Cataldo D, Munaut C, Noel A, *et al*. Matrix metalloproteinases and TIMP-1 production by peripheral blood granulocytes from COPD patients and asthmatics. *Allergy* 2001;56(2):145-51.
17. Finlay GA, O'Driscoll LR, Russell KJ, *et al*. Matrix metalloproteinase expression and production by alveolar macrophages in emphysema. *Am J Respir Crit Care Med* 1997;156(1):240-7.
18. Finlay GA, Russell KJ, McMahon KJ, *et al*. Elevated levels of matrix metalloproteinases in bronchoalveolar lavage fluid of emphysematous patients. *Thorax* 1997;52(6):502-6.
19. Hiemstra PS, van WS, Stolk J. Neutrophil serine proteinases and defensins in chronic obstructive pulmonary disease: effects on pulmonary epithelium. *Eur Respir J* 1998;12(5):1200-8.
20. Lim S, Roche N, Oliver BG, *et al*. Balance of matrix metalloprotease-9 and tissue inhibitor of metalloprotease-1 from alveolar macrophages in cigarette smokers. Regulation by interleukin-10. *Am J Respir Crit Care Med* 2000;162(4 Pt 1):1355-60.
21. Zheng T, Kang MJ, Crothers K, *et al*. Role of cathepsin S-dependent epithelial cell apoptosis in IFN-gamma-induced alveolar remodeling and pulmonary emphysema. *J Immunol* 2005;174(12):8106-15.

22. Noordhoek JA, Postma DS, Chong LL, *et al*. Different modulation of decorin production by lung fibroblasts from patients with mild and severe emphysema. *COPD: Journal of Chronic Obstructive Pulmonary Disease* 2005;2:17-25.

23. Ohnishi K, Takagi M, Kurokawa Y, *et al*. Matrix metalloproteinase-mediated extracellular matrix protein degradation in human pulmonary emphysema. *Lab Invest* 1998;78(9):1077-87.

24. van Straaten JF, Coers W, Noordhoek JA, *et al*. Proteoglycan changes in the extracellular matrix of lung tissue from patients with pulmonary emphysema. *Mod Pathol* 1999;12(7):697-705.

25. Zandvoort A, Postma DS, Jonker MR, *et al*. Altered expression of the Smad signalling pathway: implications for COPD pathogenesis. *Eur Respir J* 2006;28(3):533-41.

26. Saetta M, Baraldo S, Corbino L, *et al*. CD8+ve cells in the lungs of smokers with chronic obstructive pulmonary disease. *Am J Respir Crit Care Med* 1999; 160(2):711-7.

27. Baraldo S, Oliani KL, Turato G, *et al*. The Role of Lymphocytes in the Pathogenesis of Asthma and COPD. *Curr Med Chem* 2007;14(21):2250-6.

28. Pesci A, Balbi B, Majori M, *et al*. Inflammatory cells and mediators in bronchial lavage of patients with chronic obstructive pulmonary disease. *Eur Respir J* 1998;12(2):380-6.

29. van der Strate BW, Postma DS, Brandsma CA, *et al*. Cigarette smoke-induced emphysema: A role for the B cell? *Am J Respir Crit Care Med* 2006;173(7):751-8.

30. Hogg JC, Chu F, Utokaparch S, *et al*. The nature of small-airway obstruction in chronic obstructive pulmonary disease. *N Engl J Med* 2004 24;350(26):2645-53.

31. Balzano G, Stefanelli F, Iorio C, *et al.* Eosinophilic inflammation in stable chronic obstructive pulmonary disease. Relationship with neutrophils and airway function. *Am J Respir Crit Care Med* 1999;160(5 Pt 1):1486-92.

32. Lebowitz MD, Postma DS, Burrows B. Adverse effects of eosinophilia and smoking on the natural history of newly diagnosed chronic bronchitis. *Chest* 1995;108(1):55-61.

33. Wang X, Mensinga TT, Schouten JP, *et al*. Determinants of maximally attained level of pulmonary function. *Am J Respir Crit Care Med* 2004;169(8):941-9.

34. Demedts IK, Demoor T, Bracke KR, *et al*. Role of apoptosis in the pathogenesis of COPD and pulmonary emphysema. *Respir Res* 2006;7:53.

35. Hodge S, Hodge G, Nairn J, *et al*. Increased airway granzyme b and perforin in current and ex-smoking COPD subjects. *COPD* 2006;(34):179-87.

36. Grievink L, Smit HA, Ocke MC, *et al*. Dietary intake of antioxidant (pro)-vitamins, respiratory symptoms and pulmonary function: the MORGEN study. *Thorax* 1998 ;53(3):166-71.

37. Tabak C, Arts IC, Smit HA, *et al*. Chronic obstructive pulmonary disease and intake of catechins, flavonols, and flavones: the MORGEN Study. *Am J Respir Crit Care Med* 2001;164(1):61-4.

CHAPTER 2

Genetic epidemiology of reduced lung function

Cleo C van Diemen
H Marike Boezen

"Genetics of asthma and chronic obstructive pulmonary disease", volume 218, New York, USA, 2007

INTRODUCTION

Genetic epidemiology of respiratory disease focuses on identifying genetic determinants of disease development taking environmental factors that preclude, affect or enhance this development, into account. The respiratory disease Chronic Obstructive Pulmonary Disease (COPD) is almost fully attributable to the environmental factor smoking, with the exception of the genetically predominance of α1-antitrypsine (AAT) deficiency gene, which carriers need no further environmental smoke exposure to develop a phenotypic expression of COPD. However, AAT-deficiency is accounting for a minimal number of COPD cases worldwide (<1%), since the fast majority of COPD does not have AAT-deficiency but is attributable to cigarette smoking. The role of smoking as major determinant of COPD has risen above any doubt. However, if 'the whole world' would have been smoking, COPD would have been considered to be a genetic disease, since only a minority of smokers develops COPD. Therefore, those smokers who do develop COPD seem to be genetically more susceptible to the deleterious effects of cigarette smoke than smokers who do not develop COPD. This chapter focuses on the genetic determinants of diminished lung function, taking environmental factors that preclude, affect or enhance this development, into account, with specific emphasis on different study designs that are used in genetic epidemiology.

ENVIRONMENTAL AND GENETIC FACTORS AFFECT LUNG FUNCTION

COPD is the third cause of death worldwide and is expected to increase in prevalence in the forthcoming decades [1;2]. Reduced levels of lung function, assessed by lung function measurements, provide the most important and robust phenotypes of COPD and predict progression of the disease and overall mortality [3;4].

The level of lung function is both genetically and environmentally determined. In the past decades, the environmental determinants have been extensively studied, i.e. smoking, air pollution, occupational exposure, childhood respiratory illness, diet, and exposure to respiratory allergens. The genetic determinants have been studied less frequently and consistently, and they are more difficult to study, since COPD is a disease that is merely expressed at later ages. Therefore, the 'traditional' types of genetic studies, like those performed within families, are more difficult to perform, because the parents of individuals with COPD often already have died and the children of subjects with COPD are likely to be too young to manifest airway obstruction already.

Nevertheless, a number of genetic studies on COPD have been performed, usually including small numbers of subjects and various definitions of disease status, which makes it difficult to compare results between such studies. Because of this low inter-study-

comparability with regard to COPD disease state, it makes sense to choose a robust phenotype like level of pulmonary function, which can be compared more easily between studies. Moreover, an accelerated decline in FEV_1 is a predictor for mortality from COPD [3;4], and provides additional information compared to taking only the level of FEV_1 into account.

Environmental factors affecting lung function

The rate of decline in lung function is determined by many environmental factors. Cohort studies have shown that adult smokers experience faster decline in FEV_1 than nonsmokers [5] and that this accelerated decline returns to normal levels of aging-related decline after smoking cessation [6]. Additionally, women have a faster decline in FEV_1 in response to cigarette smoke than men [7]. The rate of FEV_1 decline is also related to outdoor and indoor air pollution levels [8;9] and occupational exposure to dust, gasses and fumes [10-12]. A poor level of lung function at a later age is not necessarily caused by excessive lung function decline, but can also be the consequence of a diminished maximal attained level of lung function, which can be due to many causes. Current and cumulative cigarette smoking, the presence of respiratory symptoms, increased numbers of blood eosinophils, and increased airway responsiveness were all found to be significant predictors of reduced level of FEV_1 in a large cohort of the general population [13]. Another study demonstrated a 10.7% reduction in FEV_1 among children with a smoking parent compared to children of a nonsmoking parent, thereby emphasizing the importance of environmental tobacco smoke already in childhood [14]. To accurately assess decline in lung function or level of lung function at an older age, it is important to keep these environmental factors in mind.

Genetic factors affecting lung function

There are many approaches to determine whether there is a genetic component contributing to a reduced level of lung function, and if such a genetic component can be identified, how big its effect is. In this chapter we will discuss the different approaches to study genetic epidemiology of reduced level of lung function, together with the major results each method has provided in elucidating the role of genetics in level of lung function.

FAMILY AND TWIN STUDIES

About 30 years ago, studies already demonstrated that there is a familial resemblance in pulmonary function. Cohen *et al.* showed in 1977 that first degree relatives of COPD patients had lower rates of forced expiration than relatives of patients with non-respiratory disease or community derived controls [15]. Many studies have addressed the familial component of COPD and found that familial factors account for up to 40-50% of

variability in cross-sectional lung function levels. For example, Astemborski *et al* showed that additive genetic variation accounted for 28% of the variation in residual FEV_1 in 108 adult families and 24% of the variation in residual FEV_1/FVC adjusted for appropriate confounders [16]. A number of studies have demonstrated that the risk to develop COPD is 2-3 fold greater for smokers who have a first degree relative with COPD compared to smokers without such a relative (reviewed by Chen [17]). Moreover, lower values of $FEF_{25-75\%}$ and $FEF_{25-75\%}$/FVC in smoking and in non-smoking first degree relatives of early onset COPD patients compared to controls suggest a genetic susceptibility to develop COPD independent of smoking [18]. Indeed, this risk may increase according to higher levels of exposure to detrimental environments such as cigarette smoke, since smoking relatives showed lower $FEF_{25-75\%}$ and $FEF_{25-75\%}$/FVC.

Some family studies have addressed the familial correlation in decline of FEV_1 and found no evidence for spousal or parent-offspring correlation for decline in FEV_1 (slope) over the years. However, there was a strong correlation in slope of FEV_1 between siblings, especially in smoking-concordant pairs [19].

Twin studies are an interesting method to study the relative importance of genetic and environmental effects on a trait, since monozygotic (MZ) twins share 100% of their DNA and dizygotic (DZ) twins only 50%. A higher degree of similarities in MZ than in DZ twins suggests a genetic influence, irrespective of environmental influences. Monozygotic twins exposed to opposite environmental influences shed a light on the relative importance of the environment e.g. cigarette smoking) on a phenotypic trait such as reduced levels of lung function. Several studies have shown that intra-pair correlations for spirometric measures between MZ twins are high and highly significant, and smaller between DZ twins [20;21], although one such a study showed no such intra-pair difference in pulmonary function between MZ or DZ twins [22]. Strong indications for genetic effects on pulmonary function come from a study by Redline and colleagues [23]. They showed that the strength of the correlations in FEV_1 and FVC decrease when relatives share less of their genotype, i.e. correlations of 0.71 for MZ twins (100% shared genotype), 0.16-0.29 for relatives with 50% shared genotype, and 0.09-0.27 for relatives with 25% shared genotype [23]. A study of lung function in 203 twin pairs aged 18-34 years detected significant genetic variation in females and males [24]. There was evidence that the heritability was lower in males (about 0.6) than females (about 0.8). An effect of smoking on lung function was detected but accounted for less than 3% of the variance. The geometry of the airways may also be related to genetic factors, since the difference in Vmax at 60% of total lung capacity (TLC) on air and on helium-oxygen is smaller between MZ than DZ twins [25]. Two other twin studies have suggested that genetic factors are important in susceptibility to airway disease in reaction to cigarette smoke. Within identical twins who are raised apart, twins concordant for cigarette smoking or non-smoking have highly similar lung function,

whereas twin pairs discordant for smoking status have larger differences in pulmonary function [26;27].

In conclusion, these family and twin studies have demonstrated that there is a familial aggregation of lung function measures, indicating that genetic factors are important for level of lung function. Because twin studies do not require the genotype and phenotype of at least 2 generations like the other types of family studies, the late onset of the disease is not a problem and twin studies are suitable for studying COPD.

SEGREGATION ANALYSIS AND HERITABILITY

Segregation analysis evaluates whether the proportion of affected and unaffected offspring in families is consistent with Mendelian expectations. It addresses the question whether a disease is caused by a major Mendelian gene or whether different minor gene effects account for the disease.

There are conflicting results for the mode of inheritance of pulmonary function. Rybicki *et al* found evidence for a major gene influencing FEV_1 in families ascertained through a proband with COPD, but not in randomly ascertained families [28]. The Humboldt family study showed a major genetic mechanism for inheritance of FVC, but environmental factors remained significant in the segregation analyses [29;30]. The National Heart, Lung, and Blood Institute (NHLBI) Family Heart study also provided evidence for major genes for FEV_1 and FEV_1/FVC [31]. In the 264 members of 26 Utah Genetic Reference pedigrees, major locus inheritance was found for the FEV_1/FVC ratio, although they could not discriminate between a dominant or recessive mode of inheritance. No evidence of major locus inheritance was found for either FEV_1 or FVC [32]. Similarly, in the Framingham study (5003 subjects form 1408 families) models with a Mendelian gene for inheritance of FEV_1 were rejected [33]. The familial correlation was greater between mothers and offspring than between fathers and offspring in models that assume no major gene. Additionally, sibling correlation exceeded parent-offspring correlation.

The Framingham study is the only study that addresses heritability of longitudinal decline in lung function. The authors describe heritability of longitudinal change in lung function over at least 15 years in two generations of the general population. Heritability factors explained a modest proportion of the population variance, but were higher when analyses were restricted to smoking status [34].

In summary, all studies indicate a significant heritability of pulmonary function, but it is not clear whether a major gene causes this effect or that several minor genes exert the effect, combined with or without environmental factors.

GENOME WIDE LINKAGE ANALYSIS

Genome wide linkage analysis is an approach to identify which chromosomal areas are involved in disease. This method compares the inheritance of disease with inheritance of genetic markers in families with multiple affected members. If the disease is co-inherited with the marker, this suggests that a disease susceptibility gene is located close to the marker of that chromosome.
The first studies on linkage analyses for COPD were performed in the Boston severe early onset COPD population. This population consists of families ascertained through probands with severe early onset COPD without AAT deficiency. Given the young age of probands, it was possible to include parents to perform these linkage analyses. Multiple chromosomes were identified to be linked to different phenotypes, e.g. pre-and postbronchodilator FEV_1 and FVC. These studies will be described in more detail in chapter 13 of this book. An overview of the genome wide linkage analyses that have been performed on lung function measures so far is given in table 1.

The Framingham cohort is based on the general population. In a genome wide screen on this cohort, FEV_1 was most influenced by the locus on chromosome 6 (LOD = 2.4), whereas chromosome 21 contained markers with the strongest linkage to FVC (LOD = 2.6) [35]. In the subsequent fine screen of the linked region, the linkage peak was narrowed to chromosome 6q27 and the LOD score for linkage of FEV_1 increased from 2.4 to 5.0 [36].

A genome-wide scan of pulmonary function measures in the NHLBI Family Heart Study identified regions on chromosomes 4 and 18 with LOD scores above 2.5, and these two chromosomes were further evaluated by incorporating additional marker genotyping. The FEV_1/FVC ratio was linked to chromosome 4 with a LOD score of 3.5. FEV_1 and FVC were suggestively linked to regions on chromosome 18 with multipoint LOD scores of 2.4 for FEV_1 and 1.5 for FVC at 31 cM (D18S843) and a LOD of 2.9 for FVC at 79 cM (D18S858) [37].

In the Utah CPH, suggestive evidence for linkage of FEV_1/FVC was found on chromosome 2 (heterogeneity LOD 2.36, dominant model) and chromosome 5 (heterogeneity LOD 2.23, recessive model). In addition, nonparametric variance component linkage analysis showed linkage of FEV_1/FVC in both of these regions, providing further support to the results [32].

A genome-wide scan for asthma phenotypes was conducted in 295 French EGEA families selected through at least one asthmatic subject. Besides linkage of asthmatic phenotypes, they also found linkage for %pred FEV_1 to chromosome 6q14, which appears to be a new region potentially linked to %pred FEV_1 [38]. Another genome wide screen within 200 families ascertained through a proband with asthma was performed by Postma *et al* [39], and in addition these authors assessed influences of early-life smoke exposure. Significant evidence for linkage of pre- and postbronchodilator FEV_1/VC was found for chromosome 2q32 (LOD 4.9, increasing to 6.03 with additional fine-mapping markers, and 3.2,

Table 1: Genome wide linkage analyses for lung function measures

Study population	Linkage	Chromosome (Cm)	LOD score	Reference
Boston early-onset COPD (n=560)	FEV_1	8p23 (2 from pter)	3.30	Palmer, 2003 [35]
	FEV_1	1p21 (136)	2.24	For every linkage, pre-BDR was also linked, but LOD score was lower
	FEV_1	8q24 (163)	2.01	
	FEV_1, FEV_1/FVC	19q13 (78)	resp. 1.94, 1.67	
	FEV_1, FEV_1/FVC	2q36 (222)	resp. 1.13, 4.42	
	FEV_1/FVC	1p31 (118)	2.52	
	FEV_1/FVC	17q21 (67)	2.44	
Boston early-onset COPD (n=585)	FEV_1/FVC	2q (222)	4.12	Silverman, 2002 [36]
	FEV_1/FVC	1(120)	1.92	
	FEV_1/FVC	17 (67)	2.03	
	FEV_1	12p (37)	2.43	
	FVC	1 (13)	2.05	
Framingham cohort, 1578 members of 330 families	FEV_1, FVC FEV_1, FVC FEV_1, FVC	4 (76) 6 (qterm) 21 (pterm)	resp. 1.6, 1.2 resp. 2.4, 1.1 resp. 1.2, 2.6	Joost, 2002 [37]
Framingham, fine screen	FEV_1	6q27 (184-190)	5.0	Wilk, 2003 [38]
NHLBI family heart study (n=1327, general population)	FEV_1/FVC FEV_1, FVC [31], FVC [79]	4 (28) 18 (31, 79)	3.5 resp. 2.4, 1.5, 2.9	Wilk, 2003 [39]
264 members of 24 Utah CEPH pedigrees	FEV_1/FVC FEV_1/FVC FEV_1/FVC FEV_1/FVC FEV_1/FVC	2 (222) 5 (295) 5 (297) 5 (299) 5 (303)	1.02 (1.55 in dom model) 2.64 1.92 1.06 (1.54 in rec model) 1.46	Malhotra, 2003 [32]
Boston early-onset COPD (n=416)	$FEF_{25-75\%}$/FVC $FEF_{25-75\%}$ $FEF_{25-75\%}$/FVC $FEF_{25-75\%}$/FVC FEV_1/FVC FEV_1	2 (216) 12 (35) 2 (221) 12 (35) 2 (229) 12 (36)	2.60 5.03 (smoke only) 4.12 (smoke only) 3.46 (smoke only) 4.13 3.26	DeMeo, 2004 [18]
295 French EGEA families, ascertained for asthma (n=726)	FEV_1%pred	6q14 (89)	2.94	Bouzigon, 2004 [40]
200 Dutch families ascertained for asthma (n=1183)	Pre- and post-BDR FEV_1/FVC Pre- and post BDR VC	2q32 (195) 5q (140)	resp. 4.9, 3.2 resp 1.8, 2.6	Postma, 2005 [41]

Abbreviations: COPD Chronic Obstructive Pulmonary Disease; FEV_1 Forced Expiratory Volume in one second; FVC Forced Vital Capacity; VC Vital Capacity; FEF Forced Expiratory Flow; NHLBI National Heart, Lung and Blood Institute; EGEA Epidemiological study on the Genetics and Environment of Asthma; BDR Bronchodilator; LOD Logarithm of the Odds; Cm Centimorgan

respectively), 35 cM proximal from linkage previously observed in families of probands with early-onset chronic obstructive pulmonary disease. Linkage existed for chromosome 5q for pre- and postbronchodilator VC (LOD 1.8 and 2.6, respectively). Results for pre- and postbronchodilator FEV_1 were less significant (LOD, 1.5 and 1.6, chromosomes 11p and 10q, respectively). Results were not affected by passive smoke exposure. Thus, there may be multiple genes on chromosome 2q that are important in determining presence and degree of airflow limitation in families ascertained for obstructive airway disease. Thus far, linkages with lung function on chromosome 2 are on different chromosomal locations in the different studies (table 1), which makes it difficult to identify a candidate gene is this area. The results from the linkage analyses are clearly inconsistent. Some regions have been replicated while most others have not. This indicates that the genetics of lung function is a multifactorial and complex matter. The logical next step after linkage analyses would be positional cloning of genes in the linked region. Unfortunately, such approaches are lacking in genetics of lung function or COPD. In contrast, in asthma, positional cloning of genes in linked regions has resulted in identification of 4 candidate genes for asthma, namely *A Disintegrin and Metalloprotease 33 (ADAM33)*, *Dipeptidyle peptidase 10 (DPP10)*, *PDH finger protein 11 (DHF11)* and *G protein coupled receptor 154 (GPRA)* [42]. The only study that examined candidate genes in a region of linkage in COPD (i.e. chromosome 2q) was performed by Celédon *et al.* [43], in which the *transforming growth factor-β_1* gene was identified as a risk factor for lower FEV_1/FVC. Hopefully, positional cloning will provide more candidate genes for reduced lung function and COPD in the future.

CASE-CONTROL STUDIES

In case-control studies, a large number of candidate genes, chosen for their biological role in pathogenesis in COPD, have been associated with lung function or COPD. However, for every studied gene there is (at least) one study which results contradict the others. Replication is difficult due to population stratification, ethnic differences, and small numbers of cases and controls that are usually included in studies. In chapter 14 (*of the original book, red. CvD*), these studies and the candidate genes that have been identified are being described in more detail.

POPULATION BASED LONGITUDINAL STUDIES ON DECLINE IN LUNG FUNCTION

As mentioned earlier, lung function decline is a risk factor for development of COPD, and the progression of asthma. Studies that evaluate genetic influences on lung function decline in combination with environmental factors are important for identifying subjects at risk for these disabling diseases.

The Lung Health Study group has performed multiple studies on genetics of lung function decline. Participants in the genetics study were selected from a larger cohort, which was designed to describe the natural history of cigarette-smoke associated COPD. They were all smokers with either no decline in FEV_1 (n=306, mean decline +14.9 ml/y) or a fast decline in FEV_1 (n=284, mean decline -153.6 ml/y) over a follow up period of 5 years. The reasoning behind selection of such groups was that widely divergent rates of decline in lung function would give robust phenotypes for detection of genes that contribute to COPD severity. In their initial study, they found associations of the α1-antitrypsin gene MZ genotype and microsomal epoxide hydrolase (His113/His139 haplotype), in combination with a family history of COPD, with rapid decline in FEV_1 [44]. Subsequent studies with this same population identified associations with rapid decline in FEV_1 of haplotypes of *IL-1β* and *IL-1 receptor antagonist* [45], haplotypes of *glutathione S-transferase* (*GST*)*-M1, -T1, -P1* [46], a SNP in *MMP1* and a haplotype of *MMP1* and *MMP12* [47], a SNP in *IL-4 receptor antagonist* [48], and a SNP in *β_2 adrenergic receptor* [49]. Since both non-decliners and fast-decliners had continuously smoked during follow up, these studies suggest that SNPs or a combination of risk alleles contribute to the deleterious effects of smoking on FEV_1. SNPs or haplotypes in other candidate genes (*tumor necrosis factor-α* (*TNF-α*), *vitamin-D binding protein* [44], *heme oxygenase-1* [46], *MMP9* [47], *IL-13, IL-13 receptor antagonist 1* [48]) were not associated with rapid decline in FEV_1. These studies have been described in more detail in Chapter 16: Association studies in COPD (*in the book 'Genetics of asthma and chronic obstructive pulmonary disease', red.*).

In a cohort of firefighters (97% males, 75% Caucasian), longitudinal decline in FEV_1 was linked to *IL-10* SNPs. A SNP at position 1668 was associated with a more rapid decline in FEV_1 [50]. In this same population, SNPs in *TNF-α*, and *IL-1β* were also associated with a more rapid decline in FEV_1, while SNPs in *TGFβ1* were not [51]. Hang *et al.* have shown that gene-environment interactions can influence decline in FEV_1 substantially [52]. They investigated whether SNPs in *microsomal epoxide hydrolase,* resulting in a slow and a fast enzyme activity, were associated with rapid decline in FEV_1 during 20 years of follow up. They were furthermore able to discriminate between non-exposed subjects and subjects exposed to airborne cotton dust and gram-negative bacteria endotoxin. Endotoxin exposure was associated with faster lung function decline among genotypes with slower enzyme activity but not among other genotypes. This indicates that SNPs can modify the association between occupational exposure and longitudinal lung function decline.

Recently, the effects of *GST* SNPs and daily fruit intake on decline in FEV_1 were studied in the Swiss SAPALDIA cohort that was followed up for 11 years [53]. Both daily fruit intake and SNPs in *GST-M1* and *GST-T1* had an effect on FEV_1 decline in men only, not in women. The protective effect of daily fruit intake was restricted to male persistent smokers. These results indicate complex gene-environment interactions with additional gender effects.

After the identification of *ADAM33* as a candidate gene for asthma [54], SNPs in this gene were also associated with excess decline in FEV_1 within asthmatics [55]. More recently, Van Diemen *et al* showed that these SNPs were associated with accelerated decline in FEV_1 and with development of COPD in the general population [56].

GENETICALLY ISOLATED POPULATIONS

Studies in isolated populations can be used in genetic epidemiology to identify candidate regions in a relatively small population because of their decreased genetic heterogeneity, increased linkage disequilibrium and more homogeneous environmental exposures. Studying isolated populations has the advantage that data collecting for phenotyping (e.g. lung function measurements) and genotyping can usually be performed in a logistically small area, with additionally low cost. The Hutterites of South Dakota are an isolated population of European descent from the early 17^{th} to 18^{th} century. Candidate loci for cardiovascular disease had a similar prevalence in this population compared to two samples of outbreed populations, indicating that data gathered in a founder population can be extrapolated to the general population [57]. However, it is possible that common variants in the general population are not present in the isolated population. Conversely, it is possible that disease alleles identified in isolated populations may not contribute significantly to susceptibility in outbreed populations. This was the case for the identification of the *cysteinyl leukotriene 2* (*CysLT2*) receptor as a candidate gene for asthma in the population of Tristan da Cunha [58]. The SNP was frequent in this population (0.13), but had only a minor effect in outbreed populations because of its low frequency (0.01-0.03) [59].

So far, no studies have been performed on COPD in isolated populations, although this might be a potentially efficient approach in identifying candidate regions, which should thereafter be explored further in general population samples.

Conclusions

Several genes have been studied in relation to reduced level of lung function and COPD. Choices about the most appropriate study design, or the population to study are usually made dependent of the goal of the study, e.g. whether one sets out to confirm a previously found association in a different population, or to search for new genes or pathways, that have not been studied before.

We have discussed several study designs used in genetic epidemiology to study determinants of reduced lung function. Overall, every type of design has its restrictions, varying e.g. from high cost to relatively low costs, and hypothesis-generating to hypothesis-testing. To our opinion, the most fruitful approach for future research seems to lie in combining the results of different types of study and study designs, and focusing on specific pathways; e.g. identifying candidate regions in isolated or patient populations, sequencing these regions and testing SNPs in candidate genes in relation to lung function decline in the general population. A logical further step lies in studying functionality of identified SNPs, as well as gene expression in tissue. Moreover, genetic epidemiology studies should warrant attention for the interactions of genes with the major environmental exposure that play a role in the development of disease, being cigarette smoking. Therefore, for future research, excellent longitudinally collected data on phenotype, as well as exposures, are as crucial as improving techniques for genotyping.

REFERENCES

1. WHO. World Health Report 2002. http://www.who.int/whr/2002 .
2. Murtagh E, Heaney L, Gingles J, *et al*. The prevalence of obstructive lung disease in a general population sample: the NICECOPD study. *Eur J Epidemiol* 2005; 20(5):443-453.
3. Ekberg-Aronsson M, Pehrsson K, Nilsson JA, *et al*. Mortality in GOLD stages of COPD and its dependence on symptoms of chronic bronchitis. *Respir Res* 2005; 6:98.
4. Sin DD, Man SF. Chronic obstructive pulmonary disease as a risk factor for cardiovascular morbidity and mortality. *Proc Am Thorac Soc* 2005; 2(1):8-11.
5. Kauffmann F, Drouet D, Lellouch J, *et al*. Twelve years spirometric changes among Paris area workers. *Int J Epidemiol* 1979; 8(3):201-212.
6. Camilli AE, Burrows B, Knudson RJ, *et al*. Longitudinal changes in forced expiratory volume in one second in adults. Effects of smoking and smoking cessation. *Am Rev Respir Dis* 1987; 135(4):794-799.
7. Xu X, Weiss ST, Rijcken B, *et al*. Smoking, changes in smoking habits, and rate of decline in FEV1: new insight into gender differences. *Eur Respir J* 1994; 7(6):1056-1061.
8. Tashkin DP, Detels R, Simmons M, *et al*. The UCLA population studies of chronic obstructive respiratory disease: XI. Impact of air pollution and smoking on annual change in forced expiratory volume in one second. *Am J Respir Crit Care Med* 1994; 149(5):1209-1217.
9. Brunekreef B, Fischer P, Remijn B, *et al*. Indoor air pollution and its effect on pulmonary function of adult non-smoking women: III. Passive smoking and pulmonary function. *Int J Epidemiol* 1985; 14(2):227-230.
10. Christiani DC, Ye TT, Zhang S, *et al*. Cotton dust and endotoxin exposure and long-term decline in lung function: results of a longitudinal study. *Am J Ind Med* 1999; 35(4):321-331.
11. Enarson DA, Vedal S, Chan-Yeung M. Rapid decline in FEV1 in grain handlers. Relation to level of dust exposure. *Am Rev Respir Dis* 1985; 132(4):814-817.
12. Kauffmann F, Drouet D, Lellouch J, *et al*. Occupational exposure and 12-year spirometric changes among Paris area workers. *Br J Ind Med* 1982; 39(3):221-232.
13. Wang X, Mensinga TT, Schouten JP, *et al*. Determinants of maximally attained level of pulmonary function. *Am J Respir Crit Care Med* 2004; 169(8):941-949.
14. Tager IB, Weiss ST, Munoz A, *et al*. Longitudinal study of the effects of maternal smoking on pulmonary function in children. *N Engl J Med* 1983; 309(12):699-703.
15. Cohen BH, Diamond EL, Graves CG, *et al*. A common familial component in lung cancer and chronic obstructive pulmonary disease. *Lancet* 1977; 2(8037):523-526.
16. Astemborski JA, Beaty TH, Cohen BH. Variance components analysis of forced expiration in families. *Am J Med Genet* 1985; 21(4):741-753.
17. Chen Y. Genetics and pulmonary medicine.10: Genetic epidemiology of pulmonary function. *Thorax* 1999; 54(9):818-824.
18. DeMeo DL, Carey VJ, Chapman HA, *et al*. Familial aggregation of FEF(25-75) and FEF(25-75)/FVC in families with severe, early onset COPD. *Thorax* 2004; 59(5):396-400.
19. Kurzius-Spencer M, Sherrill DL, Holberg CJ, *et al*. Familial correlation in the decline of forced expiratory volume in one second. *Am J Respir Crit Care Med* 2001; 164(7):1261-1265.
20. Hubert HB, Fabsitz RR, Feinleib M, *et al*. Genetic and environmental influences on pulmonary function in adult twins. *Am Rev Respir Dis* 1982; 125(4):409-415.
21. Redline S, Tishler PV, Lewitter FI, *et al*. Assessment of genetic and nongenetic influences on pulmonary function. A twin study. *Am Rev Respir Dis* 1987; 135(1):217-222.

22. Ghio AJ, Crapo RO, Elliott CG, *et al.* Heritability estimates of pulmonary function. *Chest* 1989; 96(4):743-746.

23. Redline S, Tishler PV, Rosner B, *et al.* Genotypic and phenotypic similarities in pulmonary function among family members of adult monozygotic and dizygotic twins. *Am J Epidemiol* 1989; 129(4):827-836.

24. Gibson JB, Martin NG, Oakeshott JG, *et al.* Lung function in an Australian population: contributions of polygenic factors and the Pi locus to individual differences in lung function in a sample of twins. *Ann Hum Biol* 1983; 10(6):547-556.

25. Man SF, Zamel N. Genetic influence on normal variability of maximum expiratory flow-volume curves. *J Appl Physiol* 1976; 41(6):874-877.

26. Hankins D, Drage C, Zamel N, *et al.* Pulmonary function in identical twins raised apart. *Am Rev Respir Dis* 1982; 125(1):119-121.

27. Webster PM, Lorimer EG, Man SF, *et al.* Pulmonary function in identical twins: comparison of nonsmokers and smokers. *Am Rev Respir Dis* 1979; 119(2):223-228.

28. Rybicki BA, Beaty TH, Cohen BH. Major genetic mechanisms in pulmonary function. *J Clin Epidemiol* 1990; 43(7):667-675.

29. Chen Y, Horne SL, Rennie DC, *et al.* Segregation analysis of two lung function indices in a random sample of young families: the Humboldt Family Study. *Genet Epidemiol* 1996; 13(1):35-47.

30. Chen Y, Rennie DC, Lockinger LA, *et al.* Major genetic effect on forced vital capacity: the Humboldt Family Study. *Genet Epidemiol* 1997; 14(1):63-76.

31. Wilk JB, Djousse L, Arnett DK, *et al.* Evidence for major genes influencing pulmonary function in the NHLBI family heart study. *Genet Epidemiol* 2000; 19(1):81-94.

32. Malhotra A, Peiffer AP, Ryujin DT, *et al.* Further evidence for the role of genes on chromosome 2 and chromosome 5 in the inheritance of pulmonary function. *Am J Respir Crit Care Med* 2003; 168(5):556-561.

33. Givelber RJ, Couropmitree NN, Gottlieb DJ, *et al.* Segregation analysis of pulmonary function among families in the Framingham Study. *Am J Respir Crit Care Med* 1998; 157(5 Pt 1):1445-1451.

34. Gottlieb DJ, Wilk JB, Harmon M, *et al.* Heritability of longitudinal change in lung function. The Framingham study. *Am J Respir Crit Care Med* 2001; 164(9):1655-1659.

35. Palmer LJ, Celedon JC, Chapman HA, *et al.* Genome-wide linkage analysis of bronchodilator responsiveness and post-bronchodilator spirometric phenotypes in chronic obstructive pulmonary disease, *Hum Mol Genet* 2003; 12(10): 1199-1210.

36. Silverman EK, Palmer LJ, Mosley, *et al.* Genome-wide linkage analysis of quantitative spirometric phenotypes in severe earlyonset chronicobstructive pulmonary disease, *Am J Hum Genet* 2002; 70(5):1229-1239.

37. Joost O, Wilk JB, Cupples LA, *et al.* Genetic loci influencing lung function: a genome-wide scan in the Framingham Study. *Am J Respir Crit Care Med* 2002; 165(6):795-799.

38. Wilk JB, DeStefano AL, Joost O, *et al.* Linkage and association with pulmonary function measures on chromosome 6q27 in the Framingham Heart Study. *Hum Mol Genet* 2003; 12(21):2745-2751.

39. Wilk JB, DeStefano AL, Arnett DK, *et al.* A genome-wide scan of pulmonary function measures in the National Heart, Lung, and Blood Institute Family Heart Study. *Am J Respir Crit Care Med 2003*; 167(11):1528-1533.

40. Bouzigon E, Dizier MH, Krahenbuhl C, *et al.* Clustering patterns of LOD scores for asthma-related phenotypes revealed by a genome-wide screen in 295 French EGEA families. *Hum Mol Genet* 2004; 13(24):3103-3113.

41. Postma DS, Meyers DA, Jongepier H, *et al.* Genomewide screen for pulmonary function in 200 families ascertained for asthma. *Am J Respir Crit Care Med* 2005; 172(4):446-452.

42. Kere J, Laitinen T. Positionally cloned susceptibility genes in allergy and asthma. *Curr Opin Immunol* 2004; 16(6):689-694.

43. Celedon JC, Lange C, Raby BA, *et al.* The transforming growth factor-beta1 (TGFB1) gene is associated with chronic obstructive pulmonary disease (COPD). *Hum Mol Genet* 2004; 13(15):1649-1656.

44. Sandford AJ, Chagani T, Weir TD, *et al.* Susceptibility genes for rapid decline of lung function in the lung health study. *Am J Respir Crit Care Med* 2001; 163(2):469-473.

45. Joos L, McIntyre L, Ruan J, *et al.* Association of IL-1beta and IL-1 receptor antagonist haplotypes with rate of decline in lung function in smokers. *Thorax* 2001; 56(11):863-866.

46. He JQ, Ruan J, Connett JE, *et al.* Antioxidant gene polymorphisms and susceptibility to a rapid decline in lung function in smokers. *Am J Respir Crit Care Med* 2002; 166(3):323-328.

47. Joos L, He JQ, Shepherdson MB, *et al.* The role of matrix metalloproteinase polymorphisms in the rate of decline in lung function. *Hum Mol Genet* 2002; 11(5):569-576.

48. He JQ, Connett JE, Anthonisen NR, *et al.* Polymorphisms in the IL13, IL13RA1, and IL4RA genes and rate of decline in lung function in smokers. *Am J Respir Cell Mol Biol* 2003; 28(3):379-385.

49. Joos L, Weir TD, Connett JE, *et al.* Polymorphisms in the beta2 adrenergic receptor and bronchodilator response, bronchial hyperresponsiveness, and rate of decline in lung function in smokers. *Thorax* 2003; 58(8):703-707.

50. Burgess JL, Fierro MA, Lantz RC, *et al.* Longitudinal decline in lung function: evaluation of interleukin-10 genetic polymorphisms in firefighters. *J Occup Environ Med* 2004; 46(10):1013-1022.

51. Yucesoy B, Johnson VJ, Kashon MK, *et al.* Cytokine gene polymorphisms and rate of decline in lung function in firefighters. *American Thoracic Society* 2005 , A920. 2005.

52. Hang J, Zhou W, Wang X, *et al.* Microsomal epoxide hydrolase, endotoxin, and lung function decline in cotton textile workers. *Am J Resp Crit Care Med* 2005; 171(2):165-170.

53. Imboden M, Downs SH, Senn O, *et al.* GST gene polymorphisms and fruit intake modify lung function decline: a population-based study. *American Thoracic Society* 2005 , A919. 2005.

54. Van Eerdewegh P, Little RD, Dupuis J, *et al.* Association of the ADAM33 gene with asthma and bronchial hyperresponsiveness. *Nature* 2002; 2002; 418:6896-430.

55. Jongepier H, Boezen HM, Dijkstra A, *et al.* Polymorphisms of the ADAM33 gene are associated with accelerated lung function decline in asthma. *Clin Exp Allergy* 2004; 34(5):757-760.

56. van Diemen CC, Postma DS, Vonk JM, *et al.* A disintegrin and metalloprotease 33 polymorphisms and lung function decline in the general population. *Am J Respir Crit Care Med* 2005; 172(3):329-333.

57. Newman DL, Hoffjan S, Bourgain C, *et al.* Are common disease susceptibility alleles the same in outbred and founder populations? *Eur J Hum Gen* 2004; 12(7):584-590.

58. Thompson MD, Storm van's Gravesande K, Galczenski H, *et al.* A cysteinyl leukotriene 2 receptor variant is associated with atopy in the population of Tristan da Cunha. *Pharmacogenetics* 2003; 13(10):641-649.

59. Pillai SG, Cousens DJ, Barnes AA, *et al.* A coding polymorphism in the CYSLT2 receptor with reduced affinity to LTD4 is associated with asthma. *Pharmacogenetics* 2004; 14(9):627-633

CHAPTER 3

A Disintegrin and Metalloprotease 33 polymorphisms and lung function decline in the population

Cleo C van Diemen
Dirkje S Postma
Judith M Vonk
Marcel Bruinenberg
Jan P Schouten
H Marike Boezen

Am J Respir Crit Care Med 2005;172(3):329-333

ABSTRACT

Rationale

ADAM33 (A Disintegrin and Metalloprotease 33) has been identified as a susceptibility gene for asthma and single nucleotide polymorphisms (SNPs) in this gene have been associated with excess decline of lung function in asthmatics.

Objectives

To assess whether SNPs in *ADAM33* are associated with accelerated lung function loss in the general population and with chronic obstructive pulmonary disease (COPD).

Methods

We have collected DNA from subjects of the Vlagtwedde/Vlaardingen cohort participating in the last survey in 1989/1990 after a follow up of 25 years. Information was collected every 3 years, including lung function measurements. We defined COPD as GOLD stage II or higher at the last survey. 1390 subjects from the cohort were genotyped for the following SNPs in *ADAM33*: F+1, Q-1, S_1, S_2, T_1, T_2, V_4, ST+5. Differences in prevalence of SNPs were analyzed with chi-square tests. Linear mixed effects models were used to analyze FEV_1 decline according to genotype.

Measurements and main results

In the whole population, mean adjusted decline was 18.7 and 12.7 ml/y in females and males respectively. Individuals homozygous for minor alleles of SNPs S_2 and Q-1 and heterozygous for SNP S_1 had a significantly accelerated decline in FEV_1 of respectively 4.9, 9.6 and 3.6 ml/y compared with wild type. We found a significantly higher prevalence of SNPs F+1, S_1, S_2 and T_2 in subjects with COPD.

Conclusions

We demonstrated that SNPs in *ADAM33* are associated with accelerated lung function decline in the general population. These SNPs are also a risk factor for COPD.

INTRODUCTION

ADAM33 (*A Disintegrin And Metalloprotease 33*) has been identified as an asthma susceptibility gene [1]. In Dutch, American, German and Korean asthma populations this finding has been replicated [2-5] in contrast to a Mexican population in which bronchial hyperresponsiveness was not tested [6]. Recently, we have shown that polymorphisms in the *ADAM33* gene not only play a role in asthma susceptibility, but also in its progression [7]. The S_2 polymorphism was associated with accelerated lung function decline in a Dutch asthma population followed for 23 years. It is yet unknown whether this is specific for asthma or that polymorphisms in *ADAM33* also affect lung function loss in the general population. A progressive decline in FEV_1 on a population level is partially related with asthma [8]. Additionally, lung function decline is a risk factor for the development of chronic obstructive pulmonary disease (COPD) and cardiovascular diseases [9]. Associations of polymorphisms in *ADAM33* with FEV_1 decline may therefore constitute a risk for the development of COPD as well. In this study, we have investigated the role of eight single nucleotide polymorphisms (SNPs) in the *ADAM33* gene on lung function decline in 1390 subjects of the Vlagtwedde/Vlaardingen cohort. This large population based cohort has been followed for a period up to 25 years, during which lung function measurements were performed every 3 years. We aimed to establish whether the *ADAM33* gene is associated with accelerated lung function decline in the general population, and whether it is a susceptibility gene for the development of COPD.

METHODS

Subjects

We have used data from the 2467 subjects of the Vlagtwedde/Vlaardingen cohort participating in the last survey in 1989/1990. This general population based cohort of exclusively Caucasians of Dutch descent started in 1965 and has been followed up for 25 years. The selection of the cohort has been described previously [10;11]. Surveys were performed every 3 years, in which information was collected on respiratory symptoms, smoking status, age and gender by the Dutch version of the British Medical Council standardized questionnaire [12]. A blood sample was taken and spirometry was performed. Details on pulmonary function measurements are provided in the online data supplement. In 1989/1990 neutrophil depot of spinned blood was collected and stored at -20°C. In 2003/2004 DNA was extracted from these samples with the QiaAmp® DNA Blood Mini Kit and checked for purity and concentration with the NanoDrop® ND-1000 UV-Vis Spectrophotometer. The study protocol was approved by the local university hospital's medical ethics committee and all participants gave their written informed consent.

Genotyping

We genotyped DNA samples of those subjects with more than 1500 ng isolated DNA available (N=1390). There were no differences in characteristics at the last survey between the selected and not-selected group (table 1).

Eight SNPs in *ADAM33*, previously described to be associated with asthma, airway hyperresponsiveness or excess decline in FEV_1 were genotyped: F+1 (G/A), Q-1 (C/T), S_1 (Val-Iso), S_2 (G/C), ST+5 (A/G), T_1 (Met-Thr), T_2 (Pro-Ser), V_4 (C/G). The SNP causes an amino acid change when this is indicated between brackets; otherwise, the base change is shown. The genotyping protocol is described in the online supplement; sequences of primers and probes are listed in table E1 in this supplement.

Table 1: Characteristics of subjects in the 1989/1990 survey who were and were not genotyped

	Genotyped (N=1390)	Not genotyped (N=1077)
Males, n (%)	714 (51.4)	559 (51.9)
Age in years, median (range)	52 (35-79)	54 (35-79)
Pack-years of smoking, median (range)	8.0 (0-161.4)	6.0 (0-158.4)
FEV_1%pred, mean (sd)	91.5 (14.6)	91.5 (15.9)
FEV_1/VC, mean (sd)	73.9 (8.7)	73.7 (9.1)
GOLD stage II or higher, n (%)	186 (13.4)	168 (15.6)

Abbreviations: FEV_1 Forced Expiratory Volume in 1 second; VC Vital Capacity; GOLD Global initiative for Obstructive Lung Diseases

Statistics

We identified subjects with COPD based on spirometry according to the GOLD criteria [13]. Subjects were considered to have COPD when they had an $FEV_1/VC < 70\%$ and $FEV_1 < 80\%$ predicted (GOLD stage II or higher) at the last survey. Differences in prevalence of rare alleles of SNPs between subjects with and without COPD were tested with chi-square tests. Relative risks and population attributable risks for the development of COPD were calculated for all SNPs as described in the online supplement.

Linear mixed effect models (LME) were used to investigate the effect of polymorphisms in the *ADAM33* gene on the annual decline in FEV_1 [14]. Time was defined as the time in years relative to the first FEV_1. FEV_1 measurements were included from the age of 30, since an individual's maximal achieved lung function is assumed to have been reached before that age and lung function is considered to be either in the plateau or in the decline phase [15]. Variables included in the model were age at entry, gender, pack-years of smoking and the first available FEV_1 after age 30 and their interaction with time. The outcome of the mean annual decline concerns females with age 30 when entered in the LME, a mean FEV_1 of the population, zero pack-years of smoking and with a wild type genotype. Additional information on LME analyses, Hardy Weinberg equilibrium, linkage disequilibrium,

haplotype analyses and permutation tests is described in the online supplement. Statistical analyses were performed using SPSS (version 12.0.1 for Windows), the statistical package R (version 1.9.1) [16] and Arlequin (version 2.0).

RESULTS

For all 8 SNPs in *ADAM33*, frequencies for the minor alleles were comparable to those reported previously: F+1 0.35; Q-1 0.125; S_1 0.084; S_2 0.28; ST+5 0.58; T_1 0.21; T_2 0.17; V_4 0.26. All SNPs were in Hardy Weinberg equilibrium and in significant linkage disequilibrium (LD).

In each LME, the mean adjusted annual decline for subjects with the wild type genotype for the SNPs was determined. The mean of these adjusted annual declines was 18.7 ml/yr (range 18.2-19.9). Males declined with 6 ml/yr less than females, which could be attributed to the height difference between males and females. Subjects who were homozygous for the rare alleles of Q-1 or S_2 had a significant excess decline in FEV_1 of respectively 9.6 (p=0.021) and 4.9 ml/yr (p=0.033) compared to the wild type. Heterozygous individuals for the S_1 SNP demonstrated a significant excess decline of 3.6 ml/yr (p=0.023) and a non significant excess decline of 6.4 ml/yr for homozygous individuals compared to the wild type (figure 1). None of the other SNPs were significantly associated with excess decline in FEV_1.

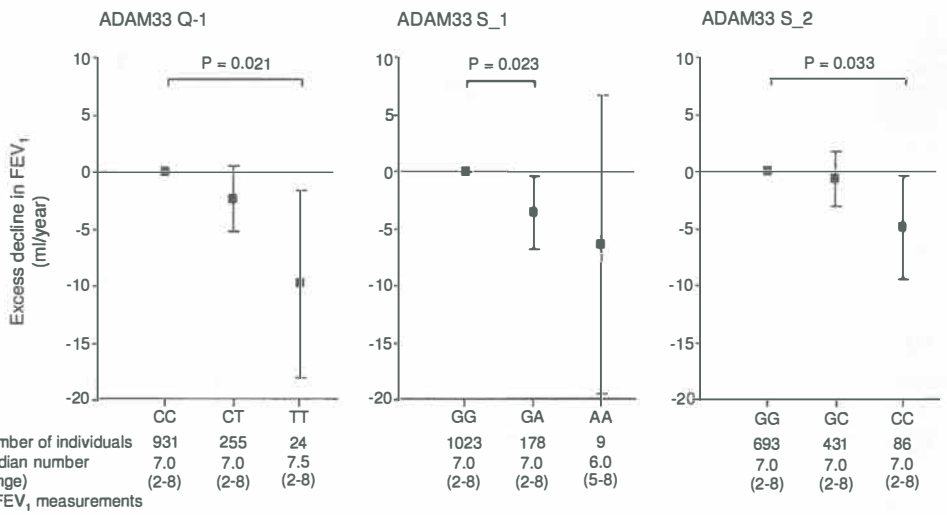


Figure 1: Excess annual decline in FEV_1 in ml/year (with 95% confidence interval) for SNPs Q-1, S_1 and S_2 compared to the wild type *Corrected for gender, age, pack-years of smoking and level of FEV_1 at age 30. Mean decline of FEV_1 for the wild type in the population is set as a reference at zero. Number of individuals and median number (range) of FEV_1 measurements are presented below the figures*

Since it would be of interest to know whether the degree of smoking might interact with SNPs on the decline in FEV_1, we performed LME models with the interaction terms pack-years*SNP in the model. These interaction-terms were not significant.

We did not find an association of level of baseline FEV_1 with any of the 8 SNPs in ADAM33, which implies that the SNPs have an effect on FEV_1 decline rather than on maximally attained lung function.

We have carried out permutation tests to assess whether our results were found due to chance. We performed 3000 permutations per SNP and ran the analysis on each of these datasets. We found the following significant p-values for the observed beta-estimates in empiric cumulative distribution: Q-1 CC genotype p=0.012, S_1 AT genotype p=0.016, S_2 TT genotype p=0.016, V_4 TT genotype p=0.038 (figure E1 in the online data supplement). Thus, for all SNPs that were associated with excess lung function decline in the true dataset, the observed beta-estimate occurred less than 5% in the empiric cumulative distribution, indicating that our results were not found due to chance. Additionally, the V_4 SNP, which was borderline significant in the original dataset (p=0.064) appeared to be significant from the permutation test.

Subjects with minor alleles for SNPs F+1, S_1, S_2 and T_2 had COPD (GOLD stage II and higher) more frequently than the remainder (P values 0.023, 0.020, 0.031, 0.039 respectively, table 2). Relative risks and population attributable risks for the development of GOLD stage II or higher are shown in the online supplement (table E2).

Table 2: Prevalence of genotypes according to COPD phenotype (GOLD stage II or higher; FEV_1/VC<70%, FEV_1 <80% predicted)

SNP		No COPD %	COPD %	P value Df=2	SNP		No COPD %	COPD %	P value Df=2
F+1	GG	46.7	38.7	**0.023**	ST+5	AA	17.5	20.0	0.630
	GA	41.4	41.9			AG	46.2	47.1	
	AA	11.9	19.4			GG	36.3	32.9	
Q-1	TT	78.2	70.9	0.095	T_1	TT	77.0	72.2	0.373
	TC	20.2	25.9			TC	21.3	25.0	
	CC	1.7	3.2			CC	1.7	2.8	
S_1	GG	85.7	77.1	**0.020**	T_2	GG	76.4	69.4	**0.039**
	GA	13.6	21.7			GA	22.1	26.5	
	AA	0.7	1.3			AA	1.5	4.1	
S_2	GG	58.3	48.1	**0.031**	V_4	CC	58.0	56.1	0.364
	GC	35.3	41.7			CG	36.5	35.5	
	CC	6.4	10.3			GG	5.5	8.4	

Abbreviations: FEV_1 Forced Expiratory Volume in 1 second; VC Vital Capacity; COPD Chronic Obstructive Pulmonary Disease; GOLD Global initiative for Obstructive Lung Diseases; df degrees of freedom; SNP Single Nucleotide Polymorphism

Haplotypes with the highest frequencies are presented in table 3. We found no significant difference in prevalence of haplotypes between subjects with and without COPD. We used haplotypes with frequencies higher than 0.01 to construct genotypes. Only the genotype with at least one minor allele for F+1, Q-1, S_1, S_2 and V_4 was associated with COPD (p=0.048). This likely reflects the individual associations found with SNPs F+1, S_1 and S_2 with COPD. In the whole population, none of the genotypes was associated with a significant excess decline in FEV_1. Thus, haplotype analysis did not provide additional information.

Table 3: Estimated frequencies of haplotypes in the whole population and in subjects with COPD
Minor alleles of the SNPs are presented; other alleles in the haplotype are wild type. Haplotypes with frequencies higher than 0.01 are shown

Haplotype	Estimated frequency in whole population	Estimated frequency in COPD
All wild type	0.303	0.266
5T+5	0.210	0.196
5T+5, V_4	0.116	0.089
F+1, 5_2, ST+5, T_1, T_2	0.106	0.128
F+1, ST+5	0.075	0.077
F+1, Q-1, 5_1, 5_2, V_4	0.069	0.101
F+1, Q-1, 5_2, 5T+5, V_4	0.023	0.023
F+1	0.019	0.021
F+1, Q-1, 5_2, 5T+5	0.012	0.017
Other	0.068	0.082

Abbreviations: COPD Chronic Obstructive Pulmonary Disease; SNP Single Nucleotide Polymorphism

DISCUSSION

We have previously shown that polymorphisms in the *ADAM33* gene are associated with accelerated decline in lung function in an asthma population [7]. Interestingly, we now present evidence that polymorphisms in *ADAM33* are associated with excess lung function decline in the general population as well. This is of great importance because it is well established that a low lung function is associated with a higher mortality risk in particular due to COPD and cardiovascular diseases [9;17-19]. Cardiovascular diseases were the most common and COPD the fifth leading cause of death in 2001 worldwide according to the WHO [20]. Since the prevalence and mortality due to COPD is expected to increase [21], this disease will become an even larger social and economic burden. COPD is a progressive disease and still poorly manageable with treatment, hence the discovery of new risk factors for this disease is important and may provide better treatments in the future [22].

Together with previously published literature, our data provides evidence that SNPs in *ADAM33* play a role in the development of asthma and COPD as well as in accelerated

lung function loss in the general population. There are no clear biological explanations for these observations. The ADAM33 protein is expressed in airway smooth muscle cells and fibroblasts and it has been proposed to contribute to the remodeling process present in asthma [1]. *ADAM33* is a member of the *ADAM* family, a family of proteins involved in cell adhesion, cell fusion, cell signaling and proteolysis [23;24]. The latter can be illustrated by the capacity to shed cytokines, growth factors or their receptors from the cell surface and the role in remodeling of extracellular matrix components. It is unknown whether the production or activity of ADAM33 in asthma is increased or reduced, yet overproduction or enhanced activity of ADAM33 may lead to excess shedding of inflammatory mediators, compatible with the enhanced airway wall inflammation present in asthma. Shedding and thereby overproduction of growth factors may furthermore induce proliferation of smooth muscle cells and fibroblasts. These features may lead to the remodeling process present in the airways of asthmatic patients. However, our data suggest that this may also reflect a general phenomenon, since polymorphisms in *ADAM33* are associated with accelerated lung function loss in the general population. Moreover, subjects carrying these polymorphisms are susceptible to COPD.

SNPs S_1 and S_2 were associated with FEV_1 decline and COPD. In contrast, F+1 and T_2 were associated with presence of COPD and not with FEV_1 decline, while SNP Q-1 was associated with FEV_1 decline and not with COPD. Our definition of COPD is based on cross-sectional data on FEV_1 and VC, which is different from FEV_1 decline. This may account for the difference in the associations found. The number of subjects with COPD may have been too small to detect the association with COPD for SNP Q-1, since we did find a trend for association (p=0.095).

In our study, three SNPs were associated with excess annual decline in FEV_1. The strongest association was found with the Q-1 SNP, which was associated with an excess decline of 9.6 ml/yr. SNP Q-1 is located in the intron before exons Q, P and R which comprise the epidermal growth factor (EGF) domain [25]. EGF signaling is important in lung morphogenesis since mice lacking the EGF receptor (EGFR) demonstrate abnormal branching and poor alveolization. Kheradmand *et al.* demonstrated that EGFR signaling regulates matrix metalloproteases (MMPs) which mediate epithelial/mesenchymal interactions during lung morphogenesis [26]. ADAM33 is closely related to MMPs but may bind EGF directly. To our knowledge, no studies so far have shown a direct binding of ADAMs to EGF or related EGFR ligands. However, some studies do provide evidence for an interaction of ADAMs with EGFR ligands. Sahin *et al*. investigated the role of ADAMs in the shedding of EGFR ligands in mouse embryonic cells derived from different ADAM knock-out mice. They identified ADAM10 as a sheddase of EGF and betacellulin and ADAM17 as a major converstase of epiregulin, TGF-α, amphiregulin and heparin-binding EGF-like growth factor (HB-EGF) [27]. ADAMs 9 and 12 were also implicated in shedding of HB-EGF. Mochizuki *et al*. found evidence of binding of ADAM28 to insulin-like growth factor

binding protein-3 with subsequent digestion [28]. It is likely that binding of these ligands to the EGF-domain of ADAMs precedes shedding by ADAMs. Therefore, if ADAM33 can bind EGF ligands to cleave them, it is also possible that it can bind them for activation. An indication for the importance of the EGF domain is that the number of mRNA transcripts containing a full EGF domain is much higher than the number of transcripts containing a full metalloprotease domain in normal airway fibroblasts [29]. However, fibroblasts from asthmatics or COPD patients may display a very different splice-pattern.

A disturbance in the EGF domain is likely to affect the regulation of ADAM33. Through alterative splicing, exon Q can be spliced out, giving rise to the β variant of ADAM33. This variant was found in 30% of *ADAM33* mRNA transcripts in pulmonary fibroblasts [29]. Since the EGF domain is incomplete, it has been suggested that the β variant prevents maturation of ADAM33 and may exert a dominant-negative effect on its protease activity [30]. The intronic SNP Q-1 may influence the splicing of the β variant [31] and thereby disturb the maturation of ADAM33. Subsequently, the protease activity may be disturbed resulting in a defect in repair of tissue after damage due to inflammation. This may lead to progressive destruction of alveolar tissue and thereby enhance accelerated decline in lung function. Our data suggest that this process may not only occur in COPD patients but also to a smaller extent in the general population depending on environmental factors.

SNPs S_1 and S_2 were associated with both excess FEV_1 decline in the general population and development of COPD. These SNPs are located in exon S, which encodes the transmembrane region. The S_1 SNP causes an amino acid change (Val to Ile), but it is unknown whether this also modifies the structure of the protein. If so, the ADAM33 protein may not be anchored correctly in the membrane and therefore not capable to exert its function. The S_2 SNP is a silent mutation and a biological explanation for the effect of this SNP is that it may be in LD with the true causative SNP. Further research on *ADAM33* should clearly lie in functional studies to elucidate the role of SNPs in *ADAM33* in lung function loss, asthma and COPD.

In conclusion, we found that polymorphisms in the *ADAM33* gene are associated with an accelerated decline in FEV_1 in the general population. In addition, we have demonstrated that these polymorphisms are not only risk factors for the development of asthma, but also for COPD. This is the first study implicating one gene to be involved in the development of both asthma and COPD. The SNPs that we have studied are all common with a frequency of at least 0.084 for the minor allele. Thus, polymorphisms in *ADAM33* constitute important risk factors for the development of respiratory diseases in a large subset of the general population.

REFERENCES

1. Van Eerdewegh P, Little RD, Dupuis J, *et al.* Association of the ADAM33 gene with asthma and bronchial hyperresponsiveness. *Nature* 2002; 418:426-430.

2. Howard TD, Postma DS, Jongepier H, *et al.* Association of a disintegrin and metalloprotease 33 (ADAM33) gene with asthma in ethnically diverse populations. *J Allergy Clin Immunol* 2003; 112:717-722.

3. Lee JH, Park HS, Park SW, *et al*. ADAM33 polymorphism: association with bronchial hyper-responsiveness in Korean asthmatics. *Clin Exp Allergy* 2004; 34:860-865.

4. Werner M, Herbon N, Gohlke H, *et al*. Asthma is associated with single-nucleotide polymorphisms in ADAM33. *Clin Exp Allergy* 2004; 34:26-31.

5. Raby BA, Silverman EK, Kwiatkowski DJ, *et al*. ADAM33 polymorphisms and phenotype associations in childhood asthma. *J Allergy Clin Immunol* 2004; 113:1071-1078.

6. Lind DL, Choudhry S, Ung N, *et al*. ADAM33 Is Not Associated with Asthma in Puerto Rican or Mexican Populations. *Am J Respir Crit Care Med* 2003; 168:1312-1316.

7. Jongepier H, Boezen HM, Dijkstra A, *et al*. Polymorphisms of the ADAM33 gene are associated with accelerated lung function decline in asthma. *Clin Exp Allergy* 2004; 34:757-760.

8. Lange P, Parner J, Vestbo J, *et al*. A 15-year follow-up study of ventilatory function in adults with asthma. *N Engl J Med* 1998;339:1194-1200.

9. Lange P, Nyboe J, Jensen G, *et al*. Ventilatory function impairment and risk of cardiovascular death and of fatal or non-fatal myocardial infarction. *Eur Respir J* 1991;4:1080-1087.

10. Rijcken B, Schouten JP, Mensinga TT, *et al*. Factors associated with bronchial responsiveness to histamine in a population sample of adults. *Am Rev Respir Dis* 1993; 147:1447-1453.

11. Van der Lende R, Kok T, Peset R, *et al*. Longterm exposure to air pollution and decline in VC and FEV1. Recent results from a longitudinal epidemiologic study in the Netherlands. *Chest* 1981; 80(1 Suppl):23-26.

12. Van der Lende R, Orie NG. The MRC-ECCS questionnaire on respiratory symptoms (use in epidemiology). *Scand J Respir Dis* 1972; 53:218-226.

13. Fabbri LM, Hurd SS, GOLD Scientific Committee, Global Strategy for the Diagnosis, Management and Prevention of COPD: 2003 update. *Eur Resp J* 2003; 22: 1-2.

14. Pinheiro JC, Bates DB. Mixed-Effects Models in S and S-Plus. New York (NY): Springer; 2000.

15. Rijcken B, Weiss ST. Longitudinal analyses of airway responsiveness and pulmonary function decline. *Am J Respir Crit Care Med* 1996; 154(6 Pt 2):S246-S249.

16. R development Core Team. R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing; 2004.

17. Friedman GD, Klatsky AL, Siegelaub AB. Lung function and risk of myocardial infarction and sudden cardiac death. *N Engl J Med* 1976; 294:1071-1075.

18. Marcus EB, Curb JD, MacLean CJ, *et al.* Pulmonary function as a predictor of coronary heart disease. *Am J Epidemiol* 1989; 129:97-104.

19. Weiss ST, Segal MR, Sparrow D, *et al*. Relation of FEV1 and peripheral blood leukocyte count to total mortality. The Normative Aging Study. *Am J Epidemiol* 1995; 142:493-498.

20. WHO. World Health Report 2002. http://www.who.int/whr/2002

21. Murray CJ, Lopez AD. Alternative projections of mortality and disability by cause 1990-2020: Global Burden of Disease Study. *Lancet* 1997; 349:1498-1504.

22. Barnes PJ, Hansel TT. Prospects for new drugs for chronic obstructive pulmonary disease. *Lancet* 2004; 364:985-996.

23. Black RA, White JM. ADAMs: focus on the protease domain. *Curr Opin Cell Biol* 1998; 10:654-659.

24. Primakoff P, Myles DG. The ADAM gene family: surface proteins with adhesion and protease activity. *Trends Genet* 2000; 16:83-87.

25. Yoshinaka T, Nishii K, Yamada K, *et al*. Identification and characterization of novel mouse and human ADAM33s with potential metalloprotease activity. *Gene* 2002; 282:227-236.

26. Kheradmand F, Rishi K, Werb Z. Signaling through the EGF receptor controls lung morphogenesis in part by regulating MT1-MMP-mediated activation of gelatinase A/MMP2. *J Cell Sci* 2002; 115(Pt 4):839-848.

27. Sahin U, Weskamp G, Kelly K, *et al*. Distinct roles for ADAM10 and ADAM17 in ectodomain shedding of six EGFR ligands. *J Cell Biol* 2004;164:769-779.

28. Mochizuki S, Shimoda M, Shiomi T, *et al*. ADAM28 is activated by MMP-7 (matrilysin-1) and cleaves insulin-like growth factor binding protein-3. *Biochem Biophys Res Commun* 2004;315:79-84.

29. Powell RM, Wicks J, Holloway JW, *et al*. The splicing and fate of ADAM33 transcripts in primary human airways fibroblasts. *Am J Respir Cell Mol Biol* 2004; 31:13-21.

30. Umland SP, Garlisi CG, Shah H, *et al*. Human ADAM33 Messenger RNA Expression Profile and Post-Transcriptional Regulation. *Am J Respir Cell Mol Biol* 2003; 29:571-582.

31. Das M, Harvey I, Chu LL, *et al*. Full-length cDNAs: more than just reaching the ends. *Physiol Genomics* 2001; 6:57-80

Online Data Supplement to paper:

A Disintegrin and Metalloprotease 33 polymorphisms and lung function decline in the population

Cleo C van Diemen, Dirkje S Postma, Judith M Vonk, Marcel Bruinenberg, Jan P Schouten, H Marike Boezen

METHODS

Spirometry

Pulmonary function measurements were performed with a water-sealed spirometer (Lode Spirograph, Lode Instruments, Groningen, The Netherlands). Measurement of inspiratory vital capacity (VC) after a deep expiration was followed by measurement of forced expiratory volume in 1 second (FEV_1). The higher of the values obtained in two technically satisfactory tracings was taken as long as the difference between the two IVC values was less than 150 ml and the difference between two FEV_1 values was less than 100 ml.

Genotyping

For all SNPs, primers and probes were obtained from Applied Biosystems TaqMan® SNP Genotyping Assays (Nieuwekerk aan de IJssel, The Netherlands), using the Assay-by-Design service for which we provided sequences. The design failed for SNP S_2, so we designed primers and probes with the Primer Express package and obtained primers from Biolegio (Malden, The Netherlands) and probes from Applied Biosystems. Sequences of all primers and probes are shown in table E1 in this online supplement. Reactions were performed in 5 µl volumes and contained 10 ng DNA, 1x Taqman Universal Mastermix (Applied Biosystems), 200 nM of each probe and 900 nM of each primer. Cycling conditions on the ABI PRISM® 7900HT Sequence Detection System (Applied Biosystems) were 10 min 95ºC followed by 40 cycles of 15 seconds 95ºC and 1 minute 60ºC. End-point fluorescence was measured immediately after cycling. Alleles were assigned using SDS 2.1 software (Applied Biosystems). Genotyping reactions failed randomly with a maximum percentage of 4.6% and a minimum of 0.6%. There was no relationship between degree of missingness and evidence of association. We regenotyped 6% of the samples and found no errors in the genotypes, therefore our data are reliable and reproducible.

Relative risks and population attributable risk percent

Relative risks (RR) for the development of COPD according to genotype were calculated in a recessive model and a dominant model. For the recessive model the RR = incidence (I) of COPD in exposed (homozygous mutant)/I in non-exposed (wild type and heterozygous individuals); for the dominant model the RR = I of COPD in heterozygotes and homozygotes/I in wild type. The population attributable risk percent (PAR%) is defined as the percent of the incidence of COPD in the population that is due to carrying a SNP. In the recessive model PAR% = (I of COPD in total population - I in wild type and heterozygotes /I in total population; in the dominant model PAR% = (I of COPD in total population - I in wild type) /I in total population.

Additional information on LME analyses

Of the 1390 subjects who were genotyped, 137 had a missing smoking history and therefore were excluded from the analyses. Data from on average 1253 subjects were included in the LME models, depending on success of genotyping. Since including the level of the first FEV_1 after age 30 and its interaction with time could introduce bias due to regression to the mean, random effects were also estimated for these variables. The results of these analyses showed no change in estimates of the variables in the model or a better fit of the model, which indicates that there was no bias due to regression-to-the-mean. Therefore, the results are presented without these random effects.

Permutation tests

We performed permutation tests to assess whether our results were found due to chance. Genotypes were randomly shuffled among individuals to produce 3000 datasets and the LME models were rerun on each of these datasets to generate a distribution of the beta-estimates under the null hypothesis. If the observed beta-estimate from the true data is found in less than 5% of the empiric cumulative distribution (p=0.05), one can assume that the observed beta-estimate is not found due to chance.

Analyses on Hardy Weinberg equilibrium, linkage equilibrium and haplotype analyses

To determine whether the SNPs were in Hardy Weinberg equilibrium and if they were in linkage disequilibrium, tests were performed with the computer program Arlequin (version 2.000). We also estimated haplotype frequencies in the whole population and in subjects with a COPD phenotype. Differences in proportions between these groups were tested with chi square tests. Haplotypes with the highest estimated frequencies in the general population were used to construct genotypes. Subjects having at least one minor allele of each SNP were selected for the genotype. With chi square tests we determined whether there was a difference in prevalence of these genotypes between subjects with

and without COPD. Also, the excess decline in FEV_1 in the whole population was determined for each constructed genotype in the LME.

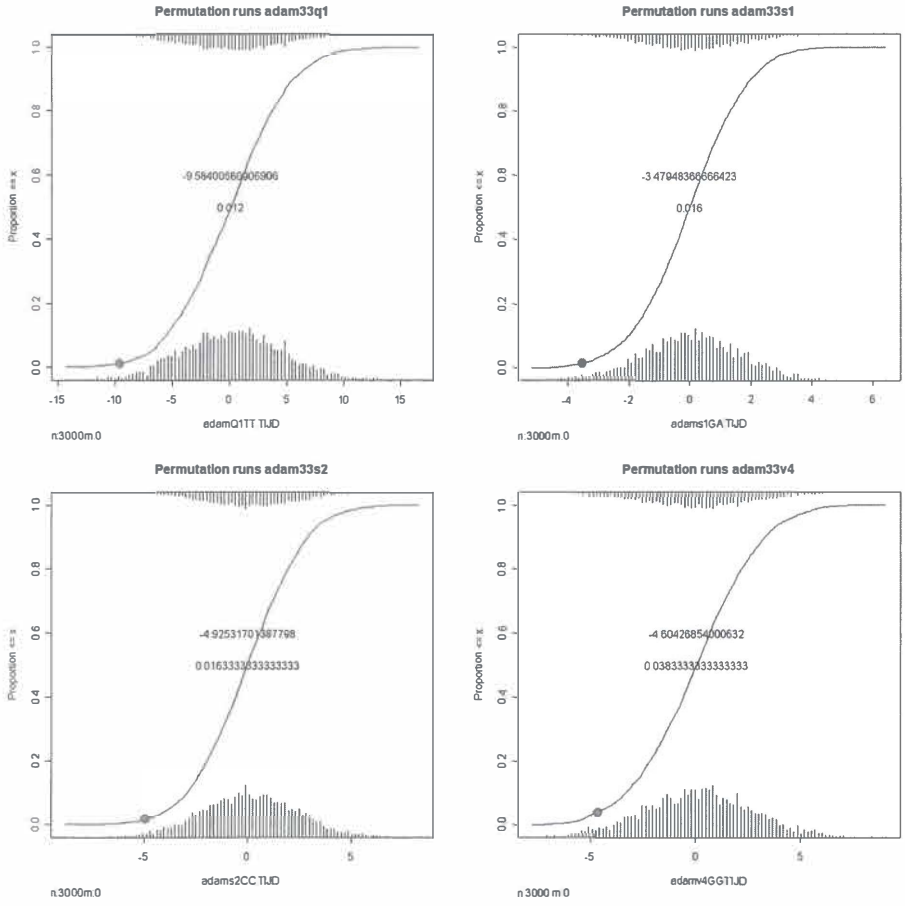


Figure E1: Permutation tests of SNPs ADAM33 Q-1, S_1, S_2 and V_4. *X-axis represents the beta estimates (decline in FEV_1 in ml/year) under the null hypothesis with the observed proportion of the cumulative distribution of the beta estimate in 3000 permutations of the genotypes on the Y-axis. The upper number in the middle of each panel is the value of the observed beta estimate of the true dataset with its according proportion in 3000 random permutations below that value.*

Table E1: Sequences of primers and probes for genotyped *ADAM33* SNPs

SNP	Primers*	Probes†
F+1	for: GGAGTGGGAATGCTGTATCTATAGC rev: GACTTCAATAAAAATACTGGGACTCGA	al1: AAGAGACAGGAATTCA al2: AGACGGGAATTCA
Q-1	for: CCCAGTGTGGACCTAGAATGG rev: GCCTGCGGGATTCAAACG	al1: CCCGACCCCTCCTT al2: ACCCGACTCCTCCTT
S_1	for: TTCCTGCTGGCCATGCT rev: CTGGGAGTCGGTAGCAACAC	al1: CCTCAGCATCCTGC al2: CCTCAGCGTCCTGC
S_2	for: CGCAGACCATGACACCTTCCT rev: GCACGCAGGGTCCCTTCT	al1: CCCAGGCGCCGGC al2: TCCCAGGGGCCGGC
ST+5	for: CAGCACATCTTTTCACTCCATACCA rev: GGTCACAGAGAACTGGGTTAAGG	al1: CAGCACCGCAGCTG al2: CAGCACCACAGCTG
T_1	for: CCCTGGTGCCTCACTCA rev: CCCAAAGATGGCCCACACA	al1: CACCCCATGGAGTTG al2: ACCCCACGGAGTTG
T_2	for: CCCTGGTGCCTCACTCA rev: CTGGGCGGCGTTCAC	al1: CCAGGACTGTCCAGTG al2: CAGGGCTGTCCAGTG
V_4	for: GGAAGGAAGGTCCCCAAAATTATGT rev: GCCCTATGGTTCGACTGAGT	al1: TCCCCTGCAGCCTG al2: TCCCCTGGAGCCTG

** for: forward primer; rev: reverse primer † bases recognizing the specific sequences are underlined; al1: allele 1, al2: allele 2*

Table E2: Prevalence of genotypes according to COPD (GOLD stage II) and relative and population attributable risks (respectively RR and PAR) for COPD in a dominant and a recessive model

SNP		No COPD %	COPD %	P value Df=2	RR dominant model	PAR% dominant model	RR recessive model	PAR% recessive model
F+1	GG	46.7	38.7	**0.023**	1.32	15.1	1.62	0.74
	GA	41.4	41.9					
	AA	11.9	19.4					
Q-1	TT	78.2	70.9	0.095	1.39	8.1	1.70	1.3
	TC	20.2	25.9					
	CC	1.7	3.2					
S_1	GG	85.7	77.1	**0.020**	1.63	8.9	1.66	0.5
	GA	13.6	21.7					
	AA	0.7	1.3					
S_2	GG	58.3	48.1	**0.031**	1.43	15.4	1.54	3.7
	GC	35.3	41.7					
	CC	6.4	10.3					
ST+5	AA	17.5	20.0	0.630	1.14	8.1	1.15	2.6
	AG	46.2	47.1					
	GG	36.3	32.9					
T_1	TT	77.0	72.2	0.373	1.30	2.2	1.54	0.7
	TC	21.3	25.0					
	CC	1.7	2.8					
T_2	GG	76.4	69.4	**0.039**	1.36	8.4	2.25	2.3
	GA	22.1	26.5					
	AA	1.5	4.1					
V_4	CC	58.0	56.1	0.364	1.07	3.0	1.46	2.9
	CG	36.5	35.5					
	GG	5.5	8.4					

CHAPTER 4

A Disintegrin And Metalloprotease 33 and Chronic Obstructive Pulmonary Disease pathophysiology

Margot ME Gosman
H Marike Boezen
Cleo C van Diemen
Jiska B Snoeck-Stroband
Thérèse S Lapperre
Pieter S Hiemstra
Nick HT ten Hacken
Jan Stolk
Dirkje S Postma
and the GLUCOLD Study Group

Thorax 2007;62(3):242-247.

ABSTRACT

Rationale

Chronic obstructive pulmonary disease (COPD) is a respiratory disorder with increasing prevalence and mortality. It is associated with airway obstruction, increased airway hyperresponsiveness (AHR), and ongoing airway and lung inflammation dominated by $CD8^+$ lymphocytes and neutrophils. Single nucleotide polymorphisms (SNPs) in A Disintegrin And Metalloprotease 33 (*ADAM33)* gene have been associated with AHR and with COPD.

Objective

To assess whether SNPs in *ADAM33* are associated with the severity of AHR and airway inflammation in COPD.

Methods

Eight SNPs in *ADAM33 (*F+1, Q-1, S_1, S_2, ST+5, T_1, T_2, V_4) were genotyped in 111 patients with COPD (96 males, 69 current smokers, mean age 62 years (SD=8), median pack-years 42 (IQR 31-55), mean postbronchodilator FEV_1% predicted 63 (SD=9)). PC_{20} methacholine, sputum, and bronchial biopsies were collected.

Results

Patients with the ST+5 AA-genotype had more severe AHR, higher numbers of sputum inflammatory cells and $CD8^+$ cells in bronchial biopsies than patients with the GG-genotype (p=0.03, p=0.05, p=0.01, respectively). $CD8^+$ cell numbers were lower in subjects carrying the minor allele of SNP T_1 and T_2, and homozygous minor variants of SNP S_2 compared to the wild-type (p=0.02, p=0.01, p=0.02, respectively).

Conclusion

This is the first study demonstrating that SNPs in a gene that confers susceptibility to COPD in the general population, i.e. *ADAM33,* are associated with AHR and airway inflammation in COPD. These findings constitute an important step forward in linking gene polymorphisms with COPD pathophysiology, thereby possibly contributing to better future treatments for this progressive and disabling disease.

INTRODUCTION

Chronic obstructive pulmonary disease (COPD) is world-wide highly prevalent and the only disease with increasing mortality rates. [1] . The disease is characterized by irreversible airflow limitation and associated with an influx of neutrophils, macrophages, and $CD8^+$ T lymphocytes in the airways [2,3] . A majority of COPD patients demonstrates airway hyperresponsiveness (AHR) [4] an exaggerated airway response to non-specific stimuli resulting in airway obstruction. The severity of AHR is positively associated with inflammation in lung tissue [5] and numbers of $CD8^+$ cells in bronchial biopsies in COPD [6]. One study reported an association of AHR with sputum inflammatory cells in COPD [7], whereas another study did not find this association [8].

Cigarette smoking is by far the most important risk factor for COPD but there is a wide range in disease severity, irrespective of the number of pack-years smoking. Furthermore, only a minority of smokers develops the disease, suggesting that besides smoking an underlying genetic constitution plays a role in the development and severity of COPD. A Disintegrin And Metalloprotease 33 (*ADAM33)* gene is a gene of putative interest for COPD. It was first identified as a susceptibility gene for asthma and AHR, [9,10] and subsequently as a susceptibility gene for COPD in a general population [11]. It furthermore is associated with accelerated lung function decline in a general [11] and asthma population [12].

ADAM33 belongs to the ADAM family of membrane-anchored metalloproteinases which play a role in a variety of processes such as cell fusion, adhesion, and signalling [13,14]. ADAM33 is expressed in both airways smooth muscle cells and lung fibroblasts. Although the overall substrates and function of *ADAM33* are yet unknown, it has been shown to be able to cleave α_2-macroglobulin [15,16], which plays a role in pulmonary defence. It is suggested that ADAM33 is involved in tissue remodelling [17], a physiologic process intricately related to airway inflammation, hyperresponsiveness and airway obstruction. The aim of this study was to test the hypothesis that SNPs in *ADAM33* are associated with the severity of both airway hyperresponsiveness and airway inflammation in sputum and bronchial biopsies of patients with COPD.

METHODS AND MATERIALS

The extended version of the methods section is available in the online data repository. One-hundred-and-fourteen patients with COPD participating in the GLUCOLD study were included [18] Patient characteristics have been described in detail previously [18]. In brief, all patients had irreversible airflow limitation and chronic respiratory symptoms, and were current or ex-smokers with at least 10 pack-years of smoking. Patients did not use a course of inhaled or oral corticosteroids within 3 months, or maintenance treatment with these

drugs within 6 months prior to randomisation. None of the patients had a history of asthma. The medical ethics committees of the Leiden University Medical Centre and the Groningen University Medical Centre approved the study, and all patients gave their written informed consent.

Pulmonary function tests have been described previously [18]. Spirometry and reversibility to salbutamol were measured. Methacholine challenge tests were performed with the 2-minute tidal breathing method [19] and expressed as the provocative concentration of methacholine causing a 20% fall in FEV_1 (PC_{20}).

Sputum induction and whole sample processing were performed according to a validated technique 20] as described in detail previously [18] Details on biopsy collection, processing and immunohistology have been published previously [21]. In brief, we collected the two best biopsies out of four paraffin embedded biopsies per patient, and used specific antibodies against T lymphocytes (CD3, CD4, and CD8), macrophages (CD68), neutrophil elastase (NE), mast cell tryptase (AA1), and eosinophils (EG2).
DNA was extracted from peripheral blood. Genotyping was performed as described previously [11] using primers and probes from Applied Biosystems TaqMan® SNP Genotyping Assays (Nieuwekerk aan de IJssel, The Netherlands).

Figure 1 shows the 8 SNPs in *ADAM33* genotyped: F+1 (G/A), Q-1 (C/T), S_1 (Val-Iso), S_2 (G/C), ST+5 (A/G), T_1 (Met-Thr), T_2 (Pro-Ser), V_4 (C/G). We based the selection of the SNPs on previous associations with airway hyperresponsiveness, excess decline in FEV_1, and/or presence of COPD [9,11,12,22,23].

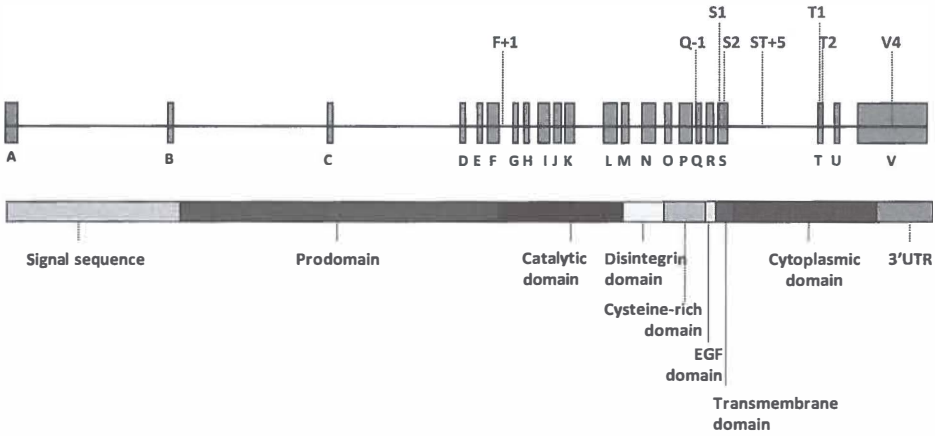


Figure 1: The exon intron structure of ADAM33 (the eight genotyped SNPs are indicated above the gene) and the domain organization of ADAM33

We used Arlequin (version 2.000) to test whether SNPs were in Hardy Weinberg equilibrium and linkage disequilibrium (LD). We investigated whether *ADAM33* SNPs are associated with PC_{20} and number and subset of inflammatory cells in sputum and

bronchial biopsies. Our primary genetic model for all SNPs was a dominant model (homozygotes and heterozygotes for the minor allele being compared as a group with homozygotes for the major allele). Additionally, SNPs with a minor allele frequency of ≥ 0.30 were entered in 1) a codominant model (three genotype groups per SNP separately) and 2) a recessive model (homozygotes and heterozygotes for the major allele being compared as a group with homozygotes for the minor allele). PC_{20} and inflammatory cells in sputum, and bronchial biopsies were log transformed to obtain a normal distribution. We performed univariate analyses using t-test and ANOVA. Multiple linear regression analyses were performed to investigate the association of polymorphisms in *ADAM33* with PC_{20}, sputum inflammatory cells, and inflammatory cells in biopsies as dependent variables. Independent variables included in the model were gender, smoking status, lung function, genotype, and the interaction of smoking status and genotype. In order to determine whether the regression estimates (with standard errors) and p values were not due to chance, we performed bootstrapping by re-sampling rows of the original data frame (using n=5000 bootstrap replicates) on the main significant outcome variables.

RESULTS

Results of univariate analyses and multiple regression analyses assuming a recessive model are shown in the online data supplement. Multiple regression analyses assuming a dominant and codominant model are presented below. The low number of subjects per haplotype did not allow haplotype analysis.

Prevalence of *ADAM33* SNPs

DNA was available from 111 out of 114 COPD patients. Clinical characteristics are presented in table 1. All genotyped SNPs were in Hardy Weinberg equilibrium and in significant LD. Table 2 shows the prevalence of the 8 SNPs. SNPs ST+5, F+1, and S_2 had a minor allele frequency ≥ 0.30, therefore these SNPs were analyzed both in a dominant model and codominant model. The frequency of SNP ST+5 in our group was 50% for each allele. We use the term AA-genotype for the wild-type and GG-genotype for the homozygous mutant genotype, as previously published [23].

Multivariate association of *ADAM33* SNPs with AHR

Dominant model Subjects with a G-allele for SNP ST+5 had a significantly higher PC_{20} compared to the AA-genotype (geometric mean (gm) 0.61 versus 0.27 mg/ml, p=0.04). SNPs F+1, Q-1, S_1, S_2, T_1, T_2, and V_4 were not significantly associated with level of AHR.

Table 1: Clinical characteristics Data are presented as mean ± standard deviation or ‡ median (25th-75th percentile).

	n = 111
Male/female	96/15
Age (years)	61.5 ± 7.7
Current smokers, n (%)	69 (62)
Smoking history (pack-years) ‡	42.0 (31.3-54.5)
FEV_1 (% pred.)	56.4 ± 9.8
FEV_1/IVC (%)	49.9 ± 8.5
Postbd. FEV_1 (% pred.)	63.2 ± 8.8
Postbd. FEV_1/IVC (%)	50.7 ± 8.5
PC_{20} methacholine (mg/ml) ‡	0.5 (0.2-2.4)

Abbreviations: FEV_1 forced expiratory volume in one second; % pred percentage of predicted value; Postbd. postbronchodilator; FEV_1/IVC forced expiratory volume in one second/inspiratory vital capacity; PC_{20} methacholinethe provocative concentration of methacholine causing a decrease in FEV_1 of 20%.

Table 2: Prevalence of *ADAM33* SNPs in patients with COPD DNA was available from 111 out of 114 COPD patients. Different numbers for the SNP genotypes (ranging from 101 to 110) are due to missing genotype data.

SNP		COPD % (n)	SNP		COPD % (n)
F+1	GG	34.9 (38)	ST+5	AA	23.7 (26)
	GA	46.8 (51)		AG	53.6 (59)
	AA	18.3 (20)		GG	22.7 (25)
Q-1	TT	67.3 (70)	T_1	TT	80.3 (86)
	TC	28.9 (30)		TC	17.8 (19)
	CC	3.8 (4)		CC	1.9 (2)
S_1	GG	72.9 (78)	T_2	GG	80.8 (88)
	GA	26.2 (28)		GA	17.4 (19)
	AA	0.9 (1)		AA	1.8 (2)
S_2	GG	44.6 (45)	V_4	CC	51.0 (55)
	GC	44.6 (45)		CG	44.4 (48)
	CC	10.8 (11)		GG	4.6 (5)

Abbreviations: SNP Single Nucleotide Polymorphism; COPD Chronic Obstructive Pulmonary Disease

Codominant model Subjects with the GG-genotype for SNP ST+5 had a significantly higher PC_{20} compared to the AA-genotype (figure 2a). SNPs F+1 and S_2 were not significantly associated with level of AHR.

Multivariate association of *ADAM33* SNPs with inflammatory cells in sputum

Dominant model Patients with one or more G-alleles of SNP ST+5 had a significantly lower total cell count in sputum than the AA-genotype (gm 176 versus $287*10^4$ cells/ml; p=0.04). Individuals with minor alleles for SNPs F+1, Q-1, S_2 had lower numbers of sputum neutrophils compared to the wild-type for those SNPs (F+1 gm 103.0 versus $134.9 *10^4$

cells/ml; p=0.05; Q-1 gm 84.5 versus 122.5*10^4 cells/ml, p=0.01; S_2 gm 94.0 versus 127.9*10^4 cells/ml, p=0.02). SNPs S_1, T_1, T_2, and V_4 were not significantly associated with the number of sputum inflammatory cells.

Codominant model Patients with the GG-genotype for SNP ST+5 had a significantly lower total cell count in sputum than the AA-genotype, and subjects with the AG-genotype tended to have a lower total cell count (figure 2b).
Heterozygous individuals for SNP S_2 had higher numbers of sputum macrophages than the wild-type (S_2 GG: gm (95% CI) 27.7*10^4/ml (20.3-37.8); GC: 39.5*10^4/ml (28.7-54.5), p=0.03; CC: 29.6*10^4/ml (17.6-49.7); p=0.80). Heterozygous individuals for SNP S_2 had also lower numbers of neutrophils (S_2 GG: gm (95% CI) 127.9*10^4/ml (97.5-168.3); GC: 93.3*10^4/ml (70.5-123.9), p=0.03; CC: 97.1*10^4/ml (61.5-153.5); p=0.24).

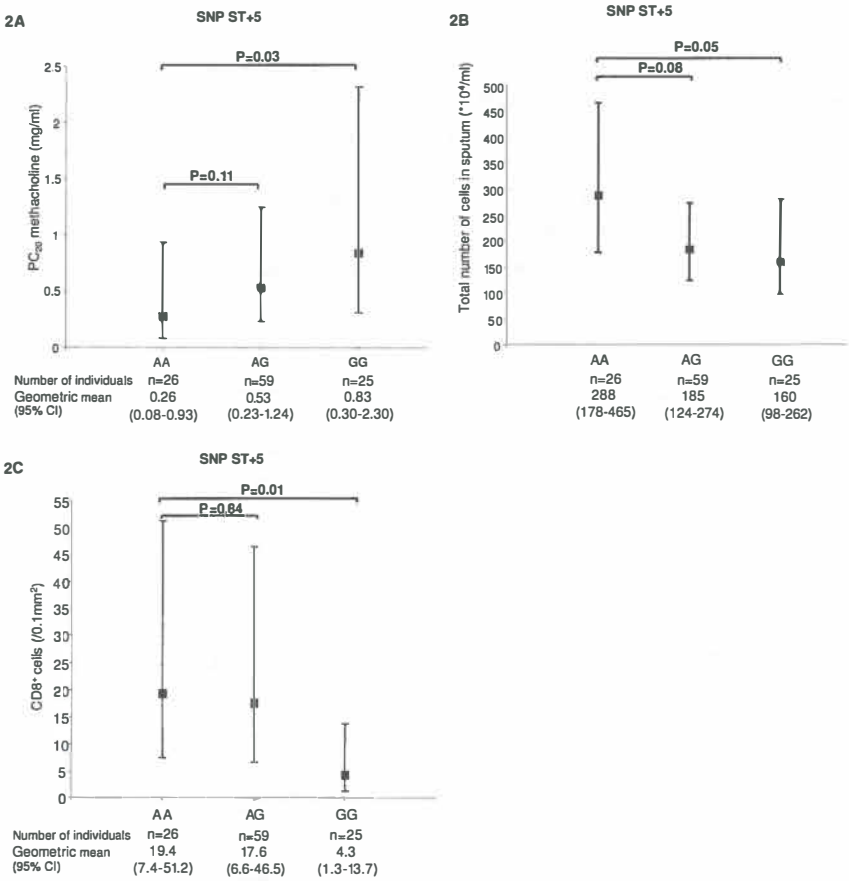


Figure 2: Association of ADAM33 SNP ST+5 with airway hyperresponsiveness (PC_{20} methacholine), sputum inflammatory cells, and $CD8^+$ cells in bronchial biopsies in patients with COPD
*A: PC_{20} methacholine per SNP ST+5 genotype B: Total sputum cell count per SNP ST+5 genotype **
C: The number of CD8+ cells in bronchial biopsies per SNP ST+5 genotype
** Total sputum cell count refers to the total of non-squamous cells in induced sputum.*
Squares represent the geometric mean; vertical bars represent the 95% confidence interval. Different numbers for the SNP genotypes are due to missing genotype data.

Multivariate association of *ADAM33* SNPs with inflammatory cells in bronchial biopsies

Dominant model Individuals with a minor allele for SNPs T_1 and T_2 had a significantly lower number of $CD8^+$ cells than the wild-type (T_1: p=0.02; T_2: p=0.01). We only present the results of SNP T_2 in figure 3a given the fact that SNPs T_1 and T_2 are in complete LD ($p= 10^{-16}$) (see online data repository; figure E1) and the association of these SNPs with the number of $CD8^+$ cells in bronchial biopsies was similar.

Codominant model Patients with the GG-genotype for SNP ST+5 had significantly lower numbers of $CD8^+$ cells in bronchial biopsies than the AA-genotype (figure 2c). Homozygous individuals for the minor allele of SNP S_2 had significantly lower numbers of $CD8^+$ cells in bronchial biopsies compared to the wild-type (figure 3b) and lower numbers of plasma cells (S_2 GG: gm (95% CI) 13.5/0.1 mm^2 (8.9-20.7); GC: 11.8/0.1 mm^2 (7.7-18.2), p=0.54; CC: 6.5/0.1 mm^2 (3.3-12.9); p=0.04). Heterozygous subjects for SNP S_2 had significantly lower numbers of $CD4^+$ cells compared with the wild-type (S_2 GG: gm (95% CI) 28.5/0.1 mm^2 (19.9-41.0); GC: 21.8/0.1 mm^2 (16.9-28.2), p=0.04; CC: 25.6/0.1 mm^2 (17.0-38.3); p=0.59). No significant associations were found with SNP F+1.

We found no significant associations between SNPs in ADAM33 and the number of mast cells or eosinophils in bronchial biopsies.
Results of bootstrapping confirmed the above presented main outcome results, with comparable and significant p-values.

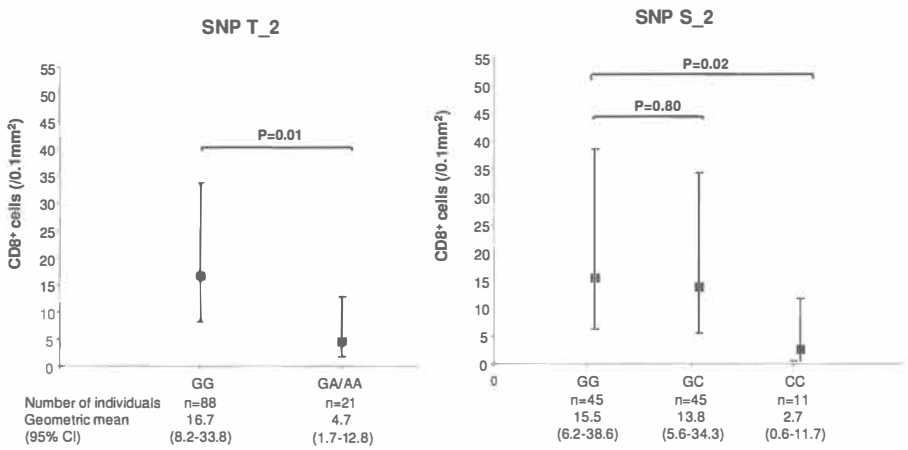


Figure 3: Association of SNPs in ADAM33 with the number of $CD8^+$ cells in bronchial biopsies in COPD
A: Per genotype for SNP T_2
B: Per genotype for SNP S_2
Squares represent the geometric mean; vertical bars represent the 95% confidence interval. Different numbers for the SNP genotypes are due to missing genotype data.

DISCUSSION

The important message of this study is that SNPs in *ADAM33* are associated with the pathophysiology of COPD. Patients with the AA-genotype for SNP ST+5 had more severe AHR, higher numbers of sputum inflammatory cells, and higher numbers of $CD8^+$ cells in bronchial biopsies than those with the GG-genotype. Moreover, individuals with the minor allele of SNP T_1 and T_2, and homozygous individuals for the minor allele of SNP S_2 had significantly lower numbers of $CD8^+$ cells in bronchial biopsies, cells relevant to the pathology of COPD.

Prevalences of the *ADAM33* SNPs found in our COPD population are comparable to those recently reported by van Diemen *et al.*[11]. We also confirm their findings that minor alleles for SNPs F+1, S_1, and S_2 are more prevalent in patients with COPD than in healthy subjects. Regarding SNP Q-1 we demonstrate a significantly higher prevalence of the minor allele in patients with COPD compared with healthy controls, whereas van Diemen *et al*. demonstrated a trend in the same direction (see online data repository). In addition, we found a higher prevalence for the A-allele in SNP ST+5 in patients with COPD. As far as we know, this is the first study suggesting that a COPD susceptibility gene is also associated with the pathophysiological process in COPD.

ADAM33 is member of the ADAM family, a group of membrane-anchored metalloproteinases which contain both a disintegrin and a metalloproteinase domain and is expressed in multiple tissues, including lung tissue [24]. Since its function has not been unravelled yet, we can only speculate about its role in COPD. It is conceivable that ADAM33 plays a role in both remodelling and inflammation by shedding growth factors, cytokines, and their receptors from the cell surface. If a SNP alters *ADAM33* function to increased protein production this may thus result in enhanced airway inflammation. Alternatively, when ADAM33 has a diminishing effect on the release of pro-inflammatory cytokines, inflammation would increase by a decrease in ADAM33 production or function. The same could be true for the role of ADAM33 in airway remodelling, in case an alteration in ADAM33 results in an increase in growth factors resulting in proliferation of airway smooth muscle cells and lung fibroblasts. Whatever the functions of ADAM33 turn out to be, our data suggests that it is associated with inflammation and AHR.

We demonstrated an association between *ADAM33* and the severity of AHR in COPD. AHR is important to COPD given its association with accelerated FEV_1 decline [25] and increased risk of COPD mortality [26]. The exact pathophysiology underlying AHR is unclear, but it is thought to result from an inflammatory process in the airways in addition to geometric changes due to airway remodelling. As discussed above, it is possible that *ADAM33* plays a role in both these processes thereby contributing to the severity of AHR. *ADAM33* SNP ST+5 is associated with the total sputum cell count in our COPD population. When we replaced this by numbers of sputum neutrophils or macrophages, we found

significant associations with SNPs S_2, F+1, and Q-1 but not with the other SNPs. This suggests that not one specific cell type was predominantly accounting for the association of total cell count with SNP ST+5. Several studies have shown an increase in the number of neutrophils and macrophages and concentrations of pro-inflammatory cytokines like IL-8 and TNF-α in induced sputum of patients with COPD [27-29]. Intuitively one would think that the severity of AHR is associated with the degree of inflammation and in fact, we previously did demonstrate an independent, positive association between severity of AHR and total sputum cell counts in these COPD patients 30]: perhaps *ADAM33* is the missing link.

O'Shaughnessy *et al*. demonstrated an increased number of neutrophils and $CD8^+$ lymphocytes in bronchial biopsies of smokers with airflow limitation, and this increase was inversely associated with level of lung function [31]. The authors therefore hypothesized that individuals with a genetic predisposition for a higher number of $CD8^+$ cells were more susceptible to a further increase in $CD8^+$ cells which might finally result in airflow limitation. Interestingly, we found an association of SNP ST+5 with both presence of COPD (see online data repository) and the number of $CD8^+$ cells within our COPD patients. *ADAM33* has previously been associated with asthma in some [9,22,23], but not all studies [32,33]. We would like to emphasise that we are confident that our COPD patients genuinely have COPD and no asthma given the fact they all had moderate to severe airflow limitation after bronchodilation, ≥10 pack-years of smoking, and no history of doctor diagnosed asthma. We selected 8 SNPs in *ADAM33* based on previous literature. One could argue that we did not genotype all known SNPs in *ADAM33*. However, we did find consistent and significant associations between the genotyped SNPs and the severity of AHR and airway inflammation in our COPD population, indicating linkage of *ADAM33* with pathophysiological features of COPD.

A potential criticism of our study is that the sample size of 111 patients with COPD is relatively small. From a practical point of view, biopsy studies are very demanding and the issue of sample size would especially be of concern in case we did not find any associations. However, we did find significant associations between *ADAM33* and features of COPD. We set out with the *a priori* hypothesis that the same SNPs in *ADAM33* previously shown to be associated with COPD are also associated with its pathophysiology. In order to investigate our hypothesis, we performed a number of analyses to assess the association of *ADAM33* with hyperresponsiveness and airway inflammation and found associations of moderate significance. This has impact on the interpretation of the results. One could raise the issue of multiple testing being responsible for the current results and that we should have adjusted for this in our analyses. We do not agree with this for a number of reasons. Firstly, the independent variables in our analyses (*e.g.* sputum total cell count and differential cell count) are mutually related, indicating that a rigid statistical procedure like *e.g.* a Bonferroni correction for multiple testing would not do justice to

their biologically linked nature. Secondly, we did not randomly test for associations between *ADAM33* and features of COPD but had a predefined hypothesis based on previous literature.

ADAM33 is a highly polymorphic gene containing at least 58 SNPs of which we investigated 8, based on previous literature [9,11,12,22,23]. We found an association of SNP ST+5 with the severity of AHR and airway inflammation in induced sputum and bronchial biopsies in COPD. The ST+5 SNP is an intron SNP between the S-exon (transmembrane region) and the T-exon (which includes a SH_3 domain and a phosporylation site). This may have functional relevance, since non-coding introns can exert their effect by influencing alternative splicing, splicing efficiency or messenger RNA turnover.

We furthermore found an association of SNP S_2 with the inflammatory cell profile in sputum and the number of inflammatory cells in bronchial biopsies. A higher number of airway wall $CD8^+$ cells is associated with more severe AHR [6], by itself a risk factor for accelerated lung function loss [25]. Van Diemen *et al* [11] showed an association of the minor allele of SNP S_2 with accelerated lung function decline in a general population and of the minor allele of SNP T_2 with presence of COPD. Yet in the present study the minor alleles of these SNPs were associated with a lower number of $CD8^+$ cells. How can we reconcile these seemingly discordant findings? SNP S_2, a silent mutation, is located in the S-exon and SNP T_2 is located in the T-exon. Both SNPs are in close proximity to SNP ST+5. The significant LD between SNPs S_2, ST+5, and T_2 may suggest that the part of *ADAM33* involved in genetic susceptibility and pathophysiology of COPD lies in the 3' region. Alternatively, it still may be either at a different location within *ADAM33* or at an adjacent gene. Our study was not designed to address this question and clearly, it needs further research.

In conclusion, our study confirms *ADAM33* as a COPD susceptibility gene and is the first to extend this observation by demonstrating an association of *ADAM33* with the severity of both airway hyperresponsiveness and airway inflammation in COPD-affected individuals. These findings constitute an important step forward in linking gene polymorphisms with COPD pathophysiology, thereby possibly contributing to better future treatments for this progressive and disabling disease.

SOURCES OF SUPPORT

Netherlands Organisation for Scientific Research (NWO), the Netherlands Asthma Foundation (NAF; 37.97.74; NAF3.2.02.51), GlaxoSmithKline (NL), Leiden University Medical Centre (LUMC), and University of Groningen (RUG)

ACKNOWLEDGEMENTS

The GLUCOLD Study Group consists of: H.F. Kauffman, D. de Reus, Dept. of Allergology; H.M. Boezen, D.F. Jansen, J.M. Vonk, Dept. of Epidemiology and Bioinformatics; M.D.W. Barentsen, M. Luinge, W. Timens, M. Zeinstra-Smit, Dept. of Pathology; A.J. Luteijn, T. van der Molen, G. ter Veen, Dept. of General Practice; M.M.E. Gosman, M.P. Farenhorst, N.H.T. ten Hacken, H.A.M. Kerstjens, M.S. van Maaren, D.S. Postma, C.A. Veltman, A. Verbokkem, I. Verhage, H.K. Vink-Klooster, Dept. of Pulmonology; University Medical Centre Groningen & University of Groningen, Groningen, The Netherlands. J.B. Snoeck-Stroband, H. Thiadens, Dept. of Public Health & Primary Care; J.K. Sont, Dept. of Medical Decision Making; I. Bajema, Dept. of Pathology; J. Gast-Strookman, P.S. Hiemstra, K. Janssen, T.S. Lapperre, K.F. Rabe, A. van Schadewijk, J. Smit-Bakker, P.J. Sterk, J. Stolk, A.C.J.A. Tiré, H. van der Veen, M.M.E. Wijffels and L.N.A. Willems, Dept. of Pulmonology; Leiden University Medical Centre, Leiden, The Netherlands; T. Mauad, University of Sao Paulo, Sao Paulo, Brazil.

REFERENCES

1. Murray CJ, Lopez AD. Alternative projections of mortality and disability by cause 1990-2020: Global Burden of Disease Study. *Lancet* 1997;349:1498-504.
2. Hogg JC, Chu F, Utokaparch S, *et al*. The nature of small-airway obstruction in chronic obstructive pulmonary disease. *N Engl J Med* 2004;350:2645-53.
3. Saetta M, Turato G, Maestrelli P, *et al*. Cellular and structural bases of chronic obstructive pulmonary disease. *Am J Respir Crit Care Med* 2001;163:1304-9.
4. Tashkin DP, Altose MD, Bleecker ER, *et al*. The lung health study: airway responsiveness to inhaled methacholine in smokers with mild to moderate airflow limitation. The Lung Health Study Research Group. *Am Rev Respir Dis* 1992;145:301-10.
5. Finkelstein R, Ma HD, Ghezzo H, *et al*. Morphometry of small airways in smokers and its relationship to emphysema type and hyperresponsiveness. *Am J Respir Crit Care Med* 1995;152:267-76.
6. Mullen JB, Wiggs BR, Wright JL, *et al*. Nonspecific airway reactivity in cigarette smokers. Relationship to airway pathology and baseline lung function. *Am Rev Respir Dis* 1986;133:120-5.
7. Rutgers SR, Timens W, Tzanakis N, *et al*. Airway inflammation and hyperresponsiveness to adenosine 5'-monophosphate in chronic obstructive pulmonary disease. *Clin Exp Allergy* 2000;30:657-62.
8. Willemse BW, ten Hacken NH, Rutgers B, *et al*. Smoking cessation improves both direct and indirect airway hyperresponsiveness in COPD. *Eur Respir J* 2004;24:391-6.
9. Van Eerdewegh P, Little RD, Dupuis J, *et al*. Association of the ADAM33 gene with asthma and bronchial hyperresponsiveness. *Nature* 2002;418:426-30.
10. Blakey J, Halapi E, Bjornsdottir US, *et al*. Contribution of ADAM33 polymorphisms to the population risk of asthma. *Thorax* 2005;60:274-6.
11. van Diemen CC, Postma DS, Vonk JM, *et al*. A disintegrin and metalloprotease 33 polymorphisms and lung function decline in the general population. *Am J Respir Crit Care Med* 2005;172:329-33.
12. Jongepier H, Boezen HM, Dijkstra A, *et al*. Polymorphisms of the ADAM33 gene are associated with accelerated lung function decline in asthma. *Clin Exp Allergy* 2004;34:757-60.
13. Black RA, White JM. ADAMs: focus on the protease domain. *Curr Opin Cell Biol* 1998;10:654-9.
14. Primakoff P, Myles DG. The ADAM gene family: surface proteins with adhesion and protease activity. *Trends Genet* 2000;16:83-7.
15. Garlisi CG, Zou J, Devito KE, *et al*. Human ADAM33: protein maturation and localization. *Biochem Biophys Res Commun* 2003;301:35-43.
16. Zou J, Zhu F, Liu J, *et al*. Catalytic activity of human ADAM33. *J Biol Chem* 2004;279:9818-30.
17. Davies DE, Wicks J, Powell RM, *et al*. Airway remodeling in asthma: new insights. *J Allergy Clin Immunol* 2003;111:215-25.
18. Lapperre TS, Snoeck-Stroband JB, Gosman MM, *et al*. Dissociation of lung function and airway inflammation in chronic obstructive pulmonary disease. *Am J Respir Crit Care Med* 2004;170:499-504.
19. Sterk PJ, Fabbri LM, Quanjer PH, *et al*. Airway responsiveness. Standardized challenge testing with pharmacological, physical and sensitizing stimuli in adults. Report Working Party Standardization of Lung Function Tests, European Community for Steel and Coal. Official Statement of the European Respiratory Society. *Eur Respir J Suppl* 1993;16:53-83.
20. in 't Veen JC, de Gouw HW, Smits HH, *et al*. Repeatability of cellular and soluble markers of inflammation in induced sputum from patients with asthma. *Eur Respir J* 1996;9:2441-7.
21. Lapperre TS, Postma DS, Gosman MM, *et al*. Relation between duration of smoking cessation and bronchial inflammation in COPD. *Thorax* 2006;61:115-21.

22. Howard TD, Postma DS, Jongepier H, *et al*. Association of a disintegrin and metalloprotease 33 (ADAM33) gene with asthma in ethnically diverse populations. *J Allergy Clin Immunol* 2003;112:717-22.

23. Werner M, Herbon N, Gohlke H, *et al*. Asthma is associated with single-nucleotide polymorphisms in ADAM33. *Clin Exp Allergy* 2004;34:26-31.

24. Umland SP, Garlisi CG, Shah H, *et al*. Human ADAM33 messenger RNA expression profile and post-transcriptional regulation. *Am J Respir Cell Mol Biol* 2003;29:571-82.

25. Tashkin DP, Altose MD, Connett JE, *et al*. Methacholine reactivity predicts changes in lung function over time in smokers with early chronic obstructive pulmonary disease. The Lung Health Study Research Group. *Am J Respir Crit Care Med* 1996;153:1802-11.

26. Hospers JJ, Postma DS, Rijcken B, *et al*. Histamine airway hyper-responsiveness and mortality from chronic obstructive pulmonary disease: a cohort study. *Lancet* 2000;356:1313-7.

27. Bhowmik A, Seemungal TA, Sapsford RJ, *et al*. Comparison of spontaneous and induced sputum for investigation of airway inflammation in chronic obstructive pulmonary disease. *Thorax* 1998;53:953-6.

28. Keatings VM, Collins PD, Scott DM, *et al*. Differences in interleukin-8 and tumor necrosis factor-alpha in induced sputum from patients with chronic obstructive pulmonary disease or asthma. *Am J Respir Crit Care Med* 1996;153:530-4.

29. Ronchi MC, Piragino C, Rosi E, *et al*. Role of sputum differential cell count in detecting airway inflammation in patients with chronic bronchial asthma or COPD. *Thorax* 1996;51:1000-4.

30. Gosman MME, van der Meulen J.M, Lapperre TS, *et al*. Airway hyperresponsiveness in COPD is more severe in females, patients with lower FEV1/VC, and higher total cell numbers in sputum. *Am J Respir Crit Care Med* 2004;167:A86.

31. O'Shaughnessy TC, Ansari TW, Barnes NC, *et al*. Inflammation in bronchial biopsies of subjects with chronic bronchitis: inverse relationship of CD8+ T lymphocytes with FEV1. *Am J Respir Crit Care Med* 1997;155:852-7.

32. Lind DL, Choudhry S, Ung N, *et al*. ADAM33 is not associated with asthma in Puerto Rican or Mexican populations. *Am J Respir Crit Care Med* 2003;168:1312-6.

33. Raby BA, Silverman EK, Kwiatkowski DJ, *et al*. ADAM33 polymorphisms and phenotype associations in childhood asthma. *J Allergy Clin Immunol* 2004;113:1071-8.

Online Data Supplement to paper:

A Disintegrin and Metalloproteinase 33 and Chronic Obstructive Pulmonary Disease Pathophysiology

Margot ME Gosman, H Marike Boezen, Cleo C van Diemen, Jiska B Snoeck-Stroband, Thérèse S Lapperre, Pieter S Hiemstra, Nick HT ten Hacken, Jan Stolk, Dirkje S Postma and the GLUCOLD Study Group

EXTENDED METHODS

Statistical analysis

We used Arlequin (version 2.000) to test whether SNPs were in Hardy Weinberg equilibrium and linkage disequilibrium (LD). Figure E1 illustrates the LD patterns with D', r and P-values in *ADAM33.* We investigated whether *ADAM33* SNPs are associated with PC_{20} and number and subset of inflammatory cells in sputum and bronchial biopsies. Our primary genetic model for all SNPs was a dominant model (homozygotes and heterozygotes for the minor allele being compared as a group with homozygotes for the major allele). Additionally, SNPs with a minor allele frequency of $\geq$ 0.30 were entered in 1) a codominant model (three genotype groups per SNP separately) and 2) a recessive model (homozygotes and heterozygotes for the major allele being compared as a group with homozygotes for the minor allele). PC_{20}, sputum inflammatory cells, and inflammatory cells in bronchial biopsies were log transformed to obtain a normal distribution. We performed univariate analyses using t-test and ANOVA. Multiple linear regression analyses were performed to investigate the association of polymorphisms in *ADAM33* with PC_{20}, sputum inflammatory cells, and inflammatory cells in biopsies as dependent variables. Independent variables included in the model were gender, smoking status, lung function, genotype, and the interaction of smoking status and genotype. We entered sputum total cell count (centered around the mean of the total population) to the model when analyzing the differential cell count in sputum, and the number of $CD3^+$ cells (centered around the mean of the total population) when analyzing $CD4^+$ and $CD8^+$ cells in bronchial biopsies.

EXTENDED RESULTS

Univariate association of ADAM33 SNPs with AHR

To provide our readership with all information, we show below the results of the dominant, codominant and recessive genetic models.

Dominant model Subjects with a G-allele for SNP ST+5 had a significantly higher PC_{20} compared to the AA-genotype (geometric mean (gm) 0.74 versus 0.27 mg/ml, p=0.02). SNPs F+1, Q-1, S_1, S_2, T_1, T_2, and V_4 were not significantly associated with AHR.

Codominant model There was a significant difference in PC_{20} methacholine between the three genotypes for SNP ST+5 (AA: gm 0.27 mg/ml; AG: 0.64 mg/ml; GG: 1.04 mg/ml; p=0.04). SNPs F+1 and S_2 were not significantly associated with AHR.

Recessive model Homozygous individuals for the minor allele of SNP S_2 had a significantly lower PC_{20} than subjects with one or more major alleles (gm 0.26 versus 0.72 mg/ml, p=0.008). SNPs F+1 and ST+5 were not significantly associated with AHR.

Univariate association of *ADAM33* SNPs with inflammatory cells in sputum

Dominant model Patients with a G-allele for SNP ST+5 had a significantly lower total sputum cell count compared to the AA-genotype (gm 121 versus $204*10^4$ cells/ml; p=0.04). SNPs F+1, Q-1, S_1, S_2, T_1, T_2, and V_4 were not significantly associated with sputum inflammatory cells.

Codominant model SNPs F+1, S_2, and ST+5 were not significantly associated with inflammatory cells in sputum.

Recessive model SNPs F+1, S_2, and ST+5 were not significantly associated with inflammatory cells in sputum.

Univariate association of *ADAM33* SNPs with inflammatory cells in bronchial biopsies

Dominant model Individuals with a minor allele for SNP S_2 had lower numbers of $CD4^+$ cells than the wild-type for that SNP (gm 41.9 versus 56.1 /0.1 mm^2, p=0.05).

Codominant model There was a significant difference between the three genotypes for SNP ST+5 and SNP S_2 in the number of $CD8^+$ cells in bronchial biopsies (gm ST+5 AA: 27.7/0.1 mm^2; AG: 21.6/0.1 mm^2; GG: 8.2/0.1 mm^2, p=0.003; S_2 GG: 21.8/0.1 mm^2; GC: 20.4/0.1 mm^2; CC: 6.2/0.1 mm^2; p=0.03). No significant associations were found with SNP F+1.

Recessive model Homozygous individuals for the G-allele of SNP ST+5 had a significantly lower number of CD8+ cells and a significantly higher number of eosinophils, compared to individuals with one or more A-alleles for SNP ST+5 (gm 8.2 versus 23.3, p=0.03 and gm 2.4 versus 0.8/0.1 mm^2, p=0.04, respectively).

Multivariate association of *ADAM33* SNPs with AHR and inflammatory cells in sputum and bronchial biopsies, assuming a recessive model

For completeness, we present below the outcome of the multivariate linear regression analyses assuming a recessive model.

SNPs F+1, S_2, and ST+5 were not significantly associated with PC_{20}, total sputum cell count, or sputum differential cell counts. The number of CD8+ cells in bronchial biopsies was significantly lower in individuals with the GG-genotype for SNP ST+5 compared with individuals with one or more A-alleles in SNP ST+5 (gm (95% CI) 7.7 (3.9-13.5) versus 18.7/0.1mm^2 (12.3-28.3), p=0.002), lower in individuals with the CC-genotype for SNP S_2 compared with individuals with one or more G-alleles in SNP S_2 (5.3 (2.4-12.0) versus 16.5/0.1 mm^2 (10.5-26.0), p= 0.007), and lower in individuals with the AA-genotype in SNP F+1 compared with individuals with one or more G-alleles in SNP F+1 (8.3 (4.4-15.4) versus 17.0/0.1 mm^2 (11.1-26.0), p=0.02). The number of eosinophils in bronchial biopsies was significantly higher in individuals with the GG-genotype for SNP ST+5 compared with individuals with one or more A-alleles in SNP ST+5 (3.5 (1.2-9.8) versus 1.2/ 0.1 mm^2 (0.6-2.5), p=0.04)

Differences in SNP prevalence between patients with COPD and controls

Previously, it has been demonstrated by van Diemen *et al.* that SNPs in *ADAM33* are associated with the presence of COPD in a general population [1]. In addition to the data presented in the current manuscript with regard to the association of *ADAM33* with airway hyperresponsiveness and airway inflammation in patients with COPD, we compared the distribution of *ADAM33* SNPs of our COPD patients with that of a population based control group.

As a control group, we selected 1097 Caucasians of Dutch descent without airflow limitation (FEV_1> 80% pred, and FEV_1/forced vital capacity> 70%) from the Vlagtwedde-Vlaardingen cohort [2,3]. Genotyping of the control group has been previously described in detail [1]. Differences in prevalence of rare alleles of SNPs between the COPD patients described in the current manuscript and controls were tested using chi-square tests. DNA was available from 1097 controls. All genotyped SNPs were in Hardy Weinberg equilibrium and in significant linkage disequilibrium. Clinical characteristics of the controls are presented in table E1.

Table E1: Clinical characteristics of the population based control group*

	Controls
Number	1097
Male/female	535/562
Age (years)	50.8 ± 9.5
Current smokers, n (%)	378 (34)
Smoking history (pack-years) ‡	6.6 (0-18.4)
FEV_1 (% pred.)	94.9 ± 11.3
FEV_1/IVC (%)	75.8 ± 6.1

*Data are presented as mean ± standard deviation or; ‡median (25th-75th percentile).
Definition of abbreviations: FEV_1 = forced expiratory volume in one second; % pred = percentage of predicted value; FEV_1/IVC = forced expiratory volume in one second/inspiratory vital capacity.

Table E2 shows that the prevalence of the minor allele of SNPs F+1, Q-1, S_1, and S_2 was significantly higher in the COPD group than the control group (p=0.04, p=0.03, p=0.003, and p=0.02, respectively), whereas the prevalence of the minor allele of SNP ST+5 was lower (p=0.02). The prevalence of SNPs T_1, T_2, and V_4 was not significantly different between both groups. With these findings, we confirm the findings by van Diemen *et al* [1] i.e. minor alleles for SNPs F+1, S_1, and S_2 are more often prevalent in patients with COPD than in subjects without airflow limitation. Regarding SNP Q-1 we found a significantly higher prevalence of the minor allele in patients with COPD compared with healthy controls, whereas van Diemen *et al.* demonstrated a trend in the same direction. In addition, we found a higher prevalence for the A-allele in SNP ST+5 in patients with COPD.

Table E2: Prevalence of genotypes in COPD patients and controls

SNP		COPD % (n)	Controls % (n)	P value Df=2	SNP		COPD % (n)	Controls % (n)	P value Df=2
F+1	GG	34.9 (38)	46.1 (491)	**0.04**	ST+5	AA	23.7 (26)	17.7 (193)	**0.02**
	GA	46.8 (51)	42.1 (455)			AG	53.6 (59)	46.9 (511)	
	AA	18.3 (20)	11.8 (128)			GG	22.7 (25)	35.4 (386)	
Q-1	TT	67.3 (70)	77.9 (844)	**0.03**	T_1	TT	80.3 (86)	77.2(795)	0.70
	TC	28.9 (30)	20.3 (220)			TC	17.8 (19)	21.2 (219)	
	CC	3.8 (4)	1.8 (19)			CC	1.9 (2)	1.6 (17)	
S_1	GG	72.9 (78)	85.3 (934)	**0.003**	T_2	GG	80.8 (88)	76.7 (817)	0.56
	GA	26.2 (28)	14.1 (154)			GA	17.4 (19)	21.8 (232)	
	AA	0.9 (1)	0.6 (7)			AA	1.8 (2)	1.5 (16)	
S_2	GG	44.6 (45)	58.5 (626)	**0.02**	V_4	CC	51.0 (55)	58.2 (631)	0.26
	GC	44.6 (45)	34.9 (374)			CG	44.4 (48)	36.4 (394)	
	CC	10.8 (11)	6.6 (71)			GG	4.6 (5)	5.4 (58)	

REFERENCES

1. Van Diemen CC, Postma DS, Vonk JM, *et al*. A disintegrin and metalloprotease 33 polymorphisms and lung function decline in the general population. *Am J Respir Crit Care Med* 2005;172:329-33.

2. Rijcken B, Schouten JP, Mensinga TT, *et al*. Factors associated with bronchial responsiveness to histamine in a population sample of adults. *Am Rev Respir Dis* 1993;147:1447-53.

3. Van der Lende R, Kok T, Peset R, *et al*. Longterm exposure to air pollution and decline in VC and FEV1. Recent results from a longitudinal epidemiologic study in the Netherlands. *Chest* 1981;80:23-6.

CHAPTER 5

ADAM33 is associated with excess lung function decline in severe COPD, irrespective of α1-antitrypsin deficiency

Cleo C van Diemen
Dirkje S Postma
W Timens
Wim van der Bij
Bea Rutgers
H Irene Heijink
Judith M Vonk
Gerard H Koëter
H Marike Boezen

Submitted for publication

ABSTRACT

Background

Forced Expiratory Volume in 1 second (FEV_1) declines with ageing in the general population, which is accelerated in COPD and asthma. Single nucleotide polymorphisms (SNPs) in *ADAM33* have been associated with airway inflammation in COPD and with accelerated FEV_1 decline in asthma and the general population. It is yet unknown whether this occurs in COPD, nor if ADAM33 is expressed in COPD lung tissue. We therefore studied *ADAM33* polymorphisms and ADAM33 expression in severe COPD.

Methods

We collected longitudinal lung function data from 79 COPD patients prior to their lung transplantation (median (IQR) age=49 (44-53) years; FEV_1=0.7 (0.5-0.85) L; 6 (3-11) FEV_1 measurements per patient; 50% male; 42 had α1-antitrypsin (AAT)-deficiency). Genotype effects on FEV_1 decline prior to transplantation were assessed using linear mixed effect models. Immunohistochemical ADAM33 expression was performed in frozen lung tissue from patients and controls.

Results

ADAM33 ST+5 and T_2 SNPs were associated with excess FEV_1 decline (ST+5: AA genotype -37.8 ml/yr vs. GG -9.5 ml/yr, p=0.04; T_2: GG -15.5 ml/yr vs. GA/AA +15.2 ml/yr, p=0.02). Stratified analyses according to AAT-status showed similar results. ADAM33 expression was demonstrated in airway smooth muscle, endothelium, mesenchyma and epithelium in all tissues.

Conclusions

ADAM33 SNPs predict excess FEV_1 decline in patients with severe COPD with and without AAT-deficiency. Combined with well-known associations of *ADAM33* SNPs with excess FEV_1 decline in asthma and the general population, we propose that *ADAM33* is not a gene exclusive for asthma or COPD, but a gene affecting FEV_1 decline irrespective of disease type.

INTRODUCTION

Chronic Obstructive Pulmonary Disease (COPD) is the fifth cause of death worldwide according to the World Health Organization and its incidence is still increasing [1]. The disease is characterized by persistent inflammation and narrowing of the airways, destruction of lung tissue, and increased mucus secretion. Although the severity of each of these conditions varies per patient, they all contribute to the accelerated decline in lung function over time in COPD patients compared with the general population [2]. Cigarette smoking is the most important risk factor for the development and progression of COPD. However, only a minority of smokers develops COPD, suggesting a genetic contribution to COPD development. One of the genes involved may be *A Disintegrin and Metalloprotease 33* (*ADAM33*), since single nucleotide polymorphisms (SNPs) in this gene have been associated with excess decline in lung function in the general population as well as with development of COPD [3]. Recently, we have shown that specific SNPs in *ADAM33* are also associated with increased inflammation in induced sputum and bronchial biopsies in a group of patients with moderate to severe COPD [4]. We therefore hypothesize that *ADAM33* SNPs may also contribute to accelerated lung function decline in patients with established COPD.

Although SNPs in *ADAM33* have been associated with asthma and COPD in multiple studies, the function of this gene is still unknown. It is highly expressed in airway fibroblasts and airway smooth muscle cells, whereas low or no expression has been observed in the airway epithelium from asthmatic patients and healthy controls [5-9]. So far, it is unknown which cells of the lung express ADAM33 in COPD. In the present study, we investigated the expression of ADAM33 in lung tissue of patients with severe COPD with and without α1-antitrypsin (AAT)-deficiency. AAT-deficiency is a genetic constitution, thought to be the main factor for development of COPD in its carriers. Excess lung function decline may develop in AAT-deficient subjects even without cigarette smoking. We hypothesized that *ADAM33* SNPs contribute to the annual decline in forced expiratory volume in 1 second (FEV_1) in severe COPD, but not in patients with AAT-deficiency, and investigated the effect of *ADAM33* SNPs on the annual decline in the FEV_1 in both patient groups.

METHODS

Subjects

We collected longitudinal lung function data from 79 patients who had lung transplantation because of severe COPD. Lung function data were collected prior to transplantation from lung function centers where patients attended regular visits to their

lung physicians. 53% of the subjects had AAT-deficiency with either the SZ or ZZ genotype. Subject characteristics are presented in table 1.

Table 1: Characteristics of patients with severe COPD before they underwent lung transplantation
Data are presented as median (interquartile range)

	Total Group (n=79)	AAT deficient (n=42)	not AAT deficient (n=37)
sex, male/female	40/39	24/18	16/21
pack-years smoking	20.0 (12.0-27.3)	15.0 (10.0-24.5)	24.0 (17.0-34.0)
age first FEV_1, years	44.8 (39.3-47.8)	42.2 (37.8-46.4)	46.1 (40.9-49.9)
age last FEV_1, years	48.9 (43.5-53.2)	46.5 (42.4-52.8)	51.4 (45.9-54.3)
first FEV_1 (L)	0.81 (0.60-1.23)	0.92 (0.64-1.39)	0.72 (0.54-1.12)
last FEV_1 (L)	0.70 (0.49-0.85)	0.72 (0.53-0.93)	0.62 (0.45-0.83)
number of FEV_1 measurements	6 (3-11)	6 (2-12)	5 (3.5-11)

Abbreviations: FEV_1: Forced Expiratory Volume in 1 second; AAT: α1-antitrypsin

DNA isolation and genotyping

Genomic DNA was extracted from frozen lung tissue using the Qiagen® DNeasy Tissue Kit (Qiagen, Hilden, Germany) and checked for purity and concentration with the NanoDrop® ND-1000 UV-Vis Spectrophotometer (Nanodrop Technologies, Wilmington, Delaware, USA). Based on previous associations with FEV_1 decline, COPD, and severity of inflammation in COPD, we selected and genotyped 8 SNPs in *ADAM33*: F+1, Q-1, S_1, S_2, ST+5, T_1, T_2, and V_4 [3;10]. The genotyping protocol and sequences of primers and probes are presented in the online supplement and table E1.

Immunohistochemistry

COPD lung tissue was obtained from 79 individuals who underwent lung transplantation and control lung tissue from tumor resection surgery. Frozen lung sections were cut at 4 µm, dried for 20 minutes and fixed in acetone (100%). Affinity purified polyclonal antibodies designated to the cytoplasmic (n=3; Sigma-Aldrich Chemie BV, Zwijndrecht, The Netherlands; Triple Point Biologics, Forest Grove, OR, USA; Everest Biotech, Oxfordshire, UK) and catalytic (n=1; Sigma-Aldrich Chemie BV) domains of ADAM33 were commercially obtained. The immunizing peptides were the same for the two cytoplasmic antibodies from Sigma-Aldrich and Triple Point Biologics. Both antibodies and the antibody designated against the catalytic domain were produced in rabbits. The third cytoplasmic antibody was a goat produced antibody and with a different immunizing peptide. Sections were stained in a Dako autostainer Universal Staining System (Dako, Copenhagen, Denmark). Endogenous peroxidase was blocked with peroxidase blocking reagent (Dako) for 5 minutes. Sections were incubated for 1 hour with polyclonal antibodies for ADAM33 in 1:200 dilutions. Subsequently, sections were incubated for 30 minutes with peroxidase conjugated rabbit-anti-goat (RAG^{po}) polyclonal (Dako). Finally, sections were incubated for 30 minutes with peroxidase conjugated goat-anti-rabbit polyclonal (GAR^{po}) (Dako).

Sections were rinsed in PBS for 5 minutes after each incubation step. Peroxidase activity was revealed by Dakocytomation 3-amino-4-ethylcarbazole Substrate-Chromogen (Dako) for 7 minutes. Sections were counterstained with Mayer's haematoxylin for 1 minute and imbedded with Kaiser's glycerin and a cover glass.

Statistics

With the statistical program R, package "genetics", we calculated whether the *ADAM33* SNPs were in Hardy Weinberg equilibrium, and whether they were in linkage disequilibrium. We used linear regression models to study SNP effects on first and last FEV_1, adjusting for sex, age, pack-years, AAT-deficiency. Linear mixed effect (LME) models were used to investigate SNP effects on annual FEV_1 decline [11]. We modeled FEV_1 decline from age 30 onwards, since an individual's maximal achieved lung function is reached before that age and lung function is considered to be either in the plateau or in the decline phase [12]. Variables in the model were sex, genotypes, AAT-deficiency, pack-years, baseline FEV_1 (centered on the population mean), and their interaction with time. Additionally, we stratified models according to AAT-deficiency status. The mean annual decline presented refers to females aged 30, with mean population baseline FEV_1, no AAT-deficiency, zero pack-years, and the wild type genotype.

Instead of performing pre- or post-hoc power analysis and correction for multiple testing, we performed permutation tests to assess whether our results were due to chance, as described previously [2]. Statistical analyses were performed with S-Plus® version 7.0, R version 1.9.1 [13], and SPSS version 14.02.

RESULTS

Association of SNPs with decline in FEV_1 in patients with severe COPD

The minor allele frequencies of the tested *ADAM33* SNPs in the group of 79 patients with severe COPD were similar to those observed in previous studies: F+1 0.39, Q-1 0.18, S_1 0.13, S_2 0.32, ST+5 0.55, T_1 0.13, T_2 0.10, V_4 0.26. Since some SNPs have a low prevalence, and our study population is small, we decided to analyze SNPs F+1, S_2, ST+5 in dominant and co-dominant models, and the other SNPs in dominant models only. All stratified analyses were performed assuming a dominant model. The SNPs F+1 and ST+5 deviated significantly from Hardy Weinberg equilibrium (p values respectively 0.009 and 0.02). However, our study population consists only of COPD patients, and Hardy Weinberg equilibrium should actually be determined in a non-diseased population. Because the allele frequencies and distribution of the alleles were similar compared to our previously published study of the Dutch general population [3], in which these SNPs were in Hardy Weinberg equilibrium, we decided to analyze F+1 and ST+5. The SNPs F+1, Q-1, S_1, S_2

and V_4 were in strong LD, and ST+5, T_1 and T_2 as well (see figure E 1 in the additional supplement).

We determined the adjusted annual FEV_1 decline for the wild type of each SNP. The mean FEV_1 decline for the wild types was 17.2 ml/yr (range 11.5 – 38.8 ml/yr), with 0.5 ml/yr extra per pack-year smoked. SNPs ST+5 and T_2 were associated with excess lung function decline (ST+5: AA genotype -37.8 ml/y vs. GG -9.5 ml/y, p=0.04; T_2: GG -15.5 ml/y vs. GA+AA +15.2 ml/y, p=0.02) (figure 1). Stratification according to AAT-deficiency status showed that SNP ST+5 was significantly associated with excess FEV_1 decline in a dominant model in the group without AAT-deficiency, but not in the group with AAT-deficiency (figure 1). The size of the estimated effect of the T_2 SNP was similar in subjects with and without AAT-deficiency, although not significantly different from the wild type (figure 1). The permutation tests confirmed the significant findings (data not shown), indicating that these results were not found due to chance.

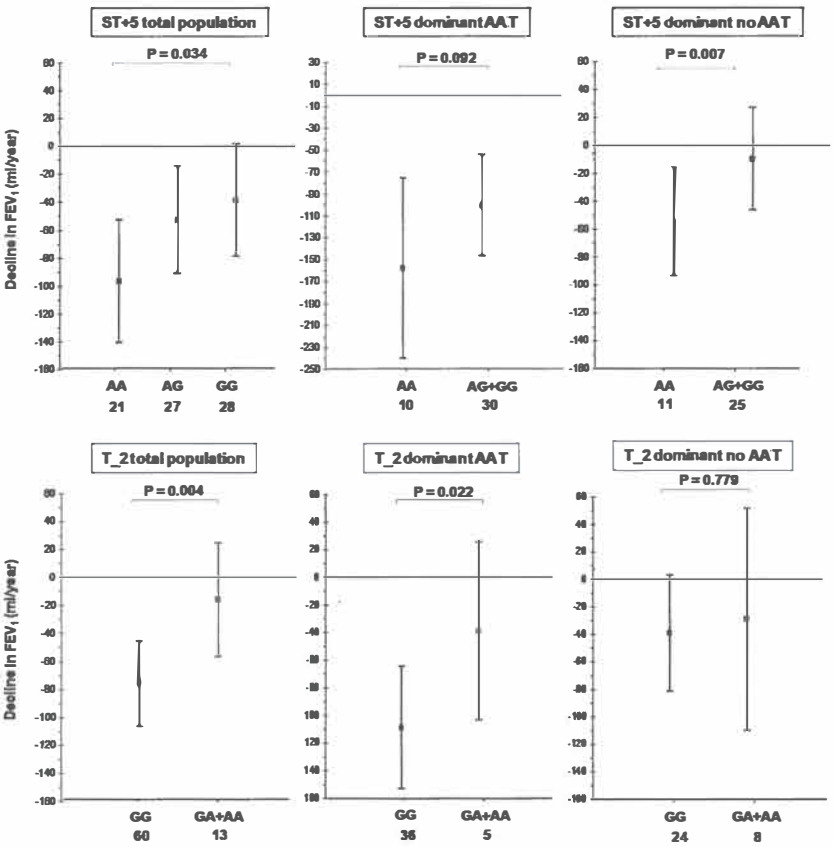


Figure 1: Association of ADAM33 SNPs ST+5 and T_2 with excess FEV_1 decline in the total population of patients with severe COPD and stratified by AAT-deficiency status

Estimates present total decline in FEV_1 per genotype, adjusted for sex, pack-years of smoking, baseline FEV_1, and AAT; error bars represent 95% confidence interval.

Abbreviations: ADAM33 A Disintegrin And Metalloprotease; SNP Single Nucleotide Polymorphism; FEV_1 Forced Expiratory Volume in 1 second; AAT α1-antitrypsin deficient

Immunohistochemistry of ADAM33 in COPD and control lung tissue

We observed ADAM33 expression in endothelium, airway smooth muscle, alveoli and inflammatory cells with all antibodies and in all tissue samples tested. In contrast to previous reports, we were also able to detect ADAM33 expression in bronchial epithelium, predominantly at the basal side (figure 2). We observed no differences in localization or intensity of expression between the groups of patients with and without AAT-deficiency, or between COPD and healthy control tissue, or with antibodies designated to the catalytic and cytoplasmic domain of ADAM33.

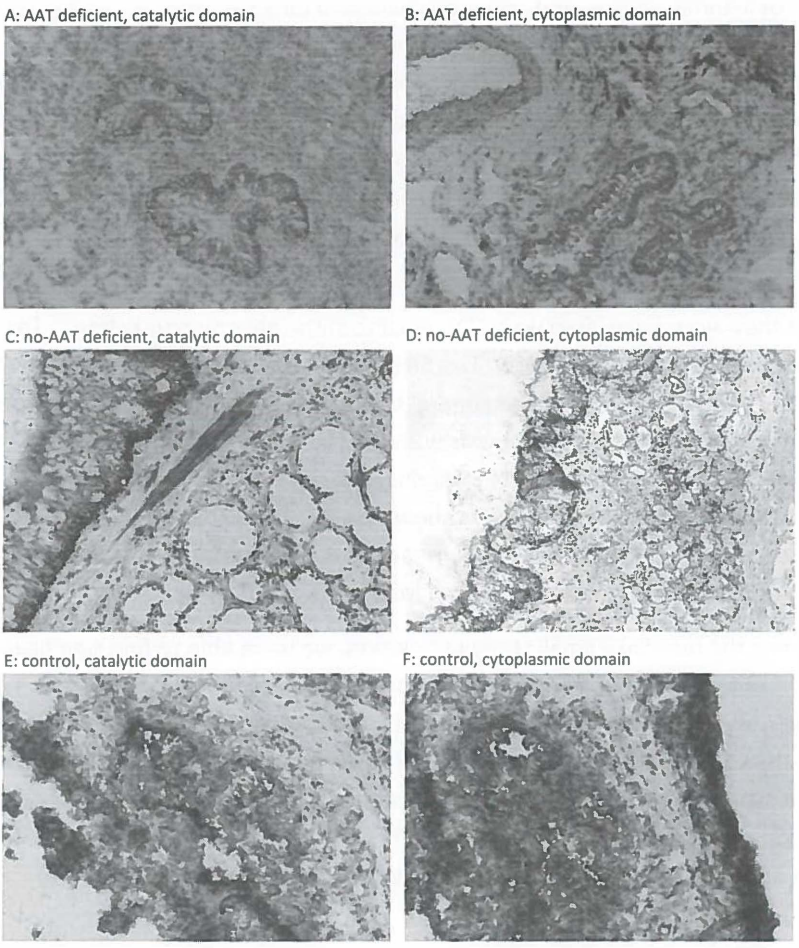


Figure 2: Immunohistochemical staining of ADAM33 in subjects with and without AAT-deficiency, and a control, with antibodies directed to the catalytic and cytoplasmic domain of ADAM33
Subject control tissue: male, current smoker, tumor lobectomy
Abbreviations: ADAM33: A Disintegrin And Metalloprotease ; AAT: α1-antitrypsin deficient

DISCUSSION

In our group of relatively young COPD patients who underwent lung transplantation, we found significant effects of *ADAM33* SNPs ST+5 and T_2 on longitudinal decline in FEV_1. We hypothesized that *ADAM33* SNPs contribute to the annual decline in FEV_1 in severe COPD, but not in patients with AAT-deficiency. However, these *ADAM33* SNPs predict additional FEV_1 decline independently from the effects already present due to AAT-deficiency. This is a remarkable finding, since it has been generally accepted that AAT-deficiency is the main underlying cause of rapid FEV_1 decline in this disease.

Previous publications have shown that genetic variations in *ADAM33* are associated with development of asthma, allergy, and impaired childhood lung function [14-20]. In the current paper, we show that *ADAM33* SNPs are associated with FEV_1 decline in patients with severe COPD. Since we previously found these SNPs to be associated with FEV_1 decline in asthmatics and the general population, we hypothesize that the *ADAM33* gene may be a gene that affects FEV_1 decline irrespective of disease. This may in addition imply that carriers of one or more *ADAM33* SNPs may develop different types of respiratory disease dependent on environmental factors and other genetic risk factors.

It is likely that SNPs Q-1, S_1 and S_2 in *ADAM33* constitute general risk factors for FEV_1 decline, since they were associated with FEV_1 decline in the general population [3]. The ST+5 SNP, however, may be specific for COPD. This SNP was only associated with excess FEV_1 decline in COPD patients, and not in the general population or asthmatics. Additionally, we previously published that within patients with mild to moderate COPD, the ST+5 SNP is associated with more severe airway hyperresponsiveness, a clinical characteristic predicting lung function loss in COPD [4]. This shows that in established COPD the ST+5 may play a role specifically in progression of the disease as demonstrated by the severity of airway hyperresponsiveness and lung function loss.

Notwithstanding the relatively small study population, we were able to find significant associations. A potential drawback of a small population size is that it may confer a type II error. However, we have performed permutation tests that confirmed our significant findings. Another potential explanation for the association of *ADAM33* SNPs with FEV_1 decline is that significant differences in a person's first or last FEV_1 cause the association. However, we found no significant associations (i.e. cross-sectionally) between the SNPs and the level of FEV_1 at the first or at the last survey in linear regression models (data not shown), indicating that this was not the case.

ADAM33 protein and mRNA expression have already been demonstrated in fibroblasts, airway smooth muscle and mesenchymal cells from asthmatics and recently to a smaller extent in epithelial cells from both asthmatics and controls [5-9;21]. However, so far it has remained unclear if and where ADAM33 is expressed in COPD lung tissue. We have

therefore performed immunohistochemical staining with antibodies against both the cytoplasmic and catalytic domain of ADAM33 on lung tissue of severe COPD patients as well as control tissue. We found similar expression profiles in smooth muscle, mesenchymal and epithelial cells in healthy and COPD lung tissue.

The function of *ADAM33* is largely unknown so far. Based on the integrin domain of *ADAM33* and the shedding capacities of *ADAM33,* we speculate that it plays a role in remodeling of the airways. SNPs in this gene may alter transcription levels, the function, folding or activity of the protein, leading to disease, together with other genetic and environmental factors [22]. Therefore, different SNPs can have different effects on ADAM33 and may thus lead to different expressions of disease. In asthma for example, ADAM33 expression increases with disease progression [21]. However, the current study did not link ADAM33 expression to specific ADAM33 SNPs. Since we did not observe differences in expression levels between COPD and healthy lung tissue, we hypothesize that ADAM33 undergoes a functional change or change in activity in COPD.

In conclusion, we found that genetic variation in *ADAM33* is associated with FEV_1 decline independent of AAT-deficiency in a unique population of patients with severe COPD followed longitudinally prior to lung transplantation. Since we previously found these SNPs to be associated with FEV_1 decline in asthmatics and the general population, we propose that the *ADAM33* gene may be a gene that affects decline of lung function irrespective of health status.

Acknowledgements

The authors thank Jan Schouten for statistical advice and Monique Lodewijk for providing immunohistochemical staining in control tissue.

Funding

This study was supported by the Dutch Asthma Foundation, grant 3.2.02.51.

REFERENCES

1. WHO. World Health Report 2002. http://www.who.int/whr/2002
2. Fletcher C, Peto R. The natural history of chronic airflow obstruction. *Br Med J* 1977; 1(6077):1645-1648.
3. van Diemen CC, Postma DS, Vonk JM, *et al.* A disintegrin and metalloprotease 33 polymorphisms and lung function decline in the general population. *Am J Respir Crit Care Med* 2005;172(3):329-333.
4. Gosman MM, Boezen HM, van Diemen CC, *et al.* A disintegrin and metalloprotease 33 and chronic obstructive pulmonary disease pathophysiology. *Thorax* 2007;62(3):242-247.
5. Umland SP, Garlisi CG, Shah H, *et al.* Human ADAM33 Messenger RNA Expression Profile and Post-Transcriptional Regulation. *Am J Respir Cell Mol Biol* 2003;29:5-582.
6. Garlisi CG, Zou J, Devito KE, *et al.* Human ADAM33: Protein maturation and localization. *Biochem Biophys Res Commun* 2003;301:1-43.
7. Powell RM, Wicks J, Holloway JW, et al. The splicing and fate of ADAM33 transcripts in primary human airways fibroblasts. *Am J Respir Cell Mol Biol* 2004;31:13-21.
8. Haitchi HM, Powell RM, Shaw TJ, *et al.* A Disintegrin and Metalloprotease 33 Expression in Asthmatic Airways and Human Embryonic Lungs. *Am J Respir Crit Care Med* 2005; 171(9):958-965
9. Lee JY, Park SW, Chang HK, *et al.* A disintegrin and metalloproteinase 33 protein in patients with asthma: relevance to airflow limitation. *Am J Respir Crit Care Med* 2006;173(7):729-735.
10. Jongepier H, Boezen HM, Dijkstra A, *et al.* Polymorphisms of the ADAM33 gene are associated with accelerated lung function decline in asthma. *Clin Exp Allergy* 2004;34(5):757-760.
11. Pinheiro JC, Bates DM. Mixed-Effects Models in S and S-Plus. New York, NY: Springer, 2000
12. Rijcken B, Weiss ST. Longitudinal analyses of airway responsiveness and pulmonary function decline. *Am J Respir Crit Care Med* 1996;154(6 Pt 2):S246-S249.
13. R development Core Team. R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing, 2004
14. Blakey J, Halapi E, Bjornsdottir US, *et al.* Contribution of ADAM33 polymorphisms to the population risk of asthma. *Thorax* 2005;60(4):274-276.
15. Cheng L, Enomoto T, Hirota T, *et al.* Polymorphisms in ADAM33 are associated with allergic rhinitis due to Japanese cedar pollen. *Clin Exp Allergy* 2004;34(8):1192-1201.
16. Kedda MA, Duffy DL, Bradley B, *et al.* ADAM33 haplotypes are associated with asthma in a large Australian population. *Eur J Hum Genet* 2006;14(9):1027-1036.
17. Noguchi E, Ohtsuki Y, Tokunaga K, *et al.* ADAM33 polymorphisms are associated with asthma susceptibility in a Japanese population. *Clin Exp Allergy* 2006;36(5):602-608.
18. Raby BA, Silverman EK, Kwiatkowski DJ, *et al.* ADAM33 polymorphisms and phenotype associations in childhood asthma. *J Allergy Clin Immunol* 2004;113(6):1071-1078.
19. Schedel M, Depner M, Schoen C, *et al.* The role of polymorphisms in ADAM33, a disintegrin and metalloprotease 33, in childhood asthma and lung function in two German populations. *Respir Res* 2006;7:91.
20. Simpson A, Maniatis N, Jury F, *et al.* Polymorphisms in A Disintegrin and Metalloprotease 33 Predict Impaired Early-Life Lung Function. *Am J Respir Crit Care Med* 2005;172(1):55-60.
21. Foley SC, Mogas AK, Olivenstein R, *et al.* Increased expression of ADAM33 and ADAM8 with disease progression in asthma. *Allergy Clin Immunol* 2007;119(4):863-871.
22. Del Mastro RG, Turenne L, Giese H, *et al.* Mechanistic role of a disease-associated genetic variant within the ADAM33 asthma susceptibility gene. *BMC Med Genet* 2007;8:46.

Online data supplement to paper:

ADAM33 is associated with excess lung function decline in severe COPD, irrespective of α1-antitrypsin deficiency

Cleo C van Diemen, Dirkje S Postma, Wim Timens, Wim van der Bij, Bea Rutgers, H Irene Heijink, Judith M Vonk, Gerard H Koëter, H Marike Boezen

METHODS

Genotyping

For all SNPs, primers and probes were obtained from Applied Biosystems TaqMan® SNP Genotyping Assays (Nieuwekerk aan de IJssel, The Netherlands), using the Assay-by-Design service for which we provided sequences. Since the design failed for SNP S_2, we designed primers and probes with the Primer Express package and obtained primers from Biolegio (Malden, The Netherlands) and probes from Applied Biosystems. Sequences of all primers and probes are shown in table E1 in this online supplement. Reactions were performed in 5 µl volumes and contained 10 ng DNA, 1x Taqman Universal Mastermix (Applied Biosystems), 200 nM of each probe and 900 nM of each primer. Cycling conditions on the ABI PRISM® 7900HT Sequence Detection System (Applied Biosystems) were 10 min 95°C followed by 40 cycles of 15 seconds 95°C and 1 minute 60°C. End-point fluorescence was measured immediately after cycling. Alleles were assigned using SDS 2.1 software (Applied Biosystems). Genotyping reactions failed randomly a maximum of 6% and a minimum of 0%. There was no relationship between degree of failure and evidence of association. We regenotyped 6% of the samples and found no errors in the genotypes, therefore our data are reliable and reproducible.

Table E1: Sequences of primers and probes for genotyped *ADAM33* SNPs

SNP	Primers	Probes
F+1	for: GGAGTGGGAATGCTGTATCTATAGC	al1: AAGAGACAGGAATTCA
	rev: GACTTCAATAAAAATACTGGGACTCGA	al2: AGACGGGAATTCA
Q-1	for: CCCAGTGTGGACCTAGAATGG	al1: CCCGACCCCTCCTT
	rev: GCCTGCGGGATTCAAACG	al2: ACCCGACTCCTCCTT
S_1	for: TTCCTGCTGGCCATGCT	al1: CCTCAGCATCCTGC
	rev: CTGGGAGTCGGTAGCAACAC	al2: CCTCAGCGTCCTGC
S_2	for: CGCAGACCATGACACCTTCCT	al1: CCCAGGCGCCGGC
	rev: GCACGCAGGGTCCCTTCT	al2: TCCCAGGGGCCGGC
ST+5	for: CAGCACATCTTTTCACTCCATACCA	al1: CAGCACCGCAGCTG
	rev: GGTCACAGAGAACTGGGTTAAGG	al2: CAGCACCACAGCTG
T_1	for: CCCTGGTGCCTCACTCA	al1: CACCCCATGGAGTTG
	rev: CCCAAAGATGGCCCACACA	al2: ACCCCACGGAGTTG
T_2	for: CCCTGGTGCCTCACTCA	al1: CCAGGACTGTCCAGTG
	rev: CTGGGCGGCGTTCAC	al2: CAGGGCTGTCCAGTG
V_4	for: GGAAGGAAGGTCCCCAAAATTATGT	al1: TCCCCTGCAGCCTG
	rev: GCCCTATGGTTCGACTGAGT	al2: TCCCCTGGAGCCTG

** for: forward primer; rev: reverse primer † bases recognizing the specific sequences are underlined; al1: allele 1, al2: allele 2*

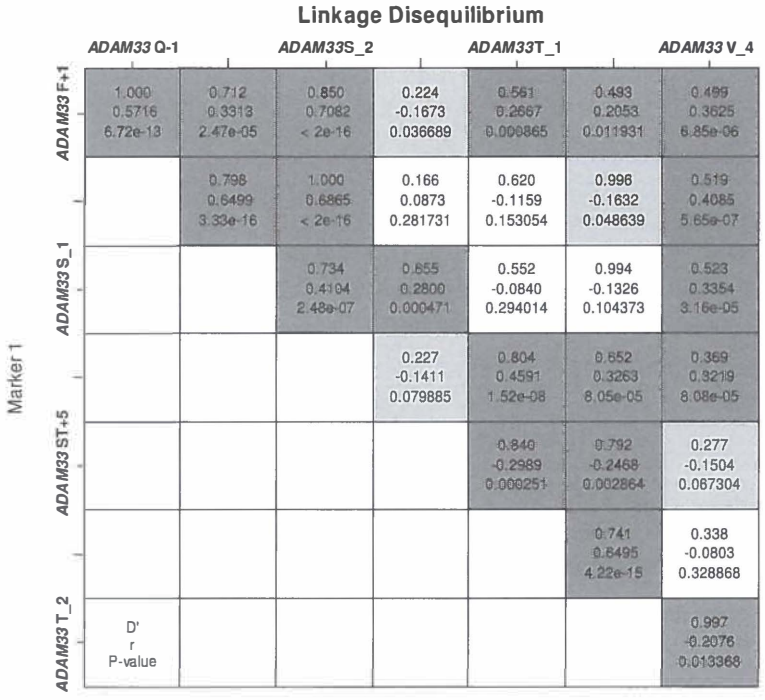


Figure E1: Linkage disequilibrium map of ADAM33 SNPs in the group of patients with severe COPD

CHAPTER 6

Genetic variation in *TIMP1* but not in *MMPs* predict excess FEV_1 decline in two populations

Cleo C van Diemen
Dirkje S Postma
Mateusz Siedliński
Henriëtte A Smit
Anneke Blokstra
H Marike Boezen

Submitted for publication

ABSTRACT

Background

An imbalance in Matrix MetalloProteases (MMPs) and Tissue Inhibitors of MMPs (TIMPs) contributes to Chronic Obstructive Pulmonary Disease (COPD) development. Prospective studies investigating Single Nucleotide Polymorphisms (SNPs) in *MMPs* and *TIMPs* with respect to COPD development and lung function decline in the general population are lacking.

Methods

We genotyped SNPs in *MMP1* (G-1607GG), *MMP2* (-1306 C/T), *MMP9* (3 tagging SNPs), *MMP12* (A-82G and Asn357Ser) and *TIMP1* (Phe124Phe and Ile158Ile) in 1390 Caucasians with multiple FEV_1 measurements from the general population. FEV_1 decline was analyzed using linear mixed effect models adjusted for confounders. Analyses of the X-chromosomal *TIMP1* gene were stratified according to sex. All significant associations were repeated in an independent general population cohort (n=1152).

Results

MMP2 -1306 TT genotype carriers had excess FEV_1 decline (-4.0 ml/yr, p=0.03) compared to wild type carriers. *TIMP1* Ile158Ile predicted significant excess FEV_1 decline in males and females. *TIMP1* Phe124Phe predicted significant excess FEV_1 decline in males only, which was replicated in the second cohort. The *MMP2* and *TIMP1* Ile158Ile associations were not replicated.

Discussion

We for the first time show that *TIMP1* Phe124Phe contributes to excess FEV_1 decline in two independent prospective cohorts. *MMPs* evidently do not contribute to FEV_1 decline in the general population.

INTRODUCTION

Chronic Obstructive Pulmonary Disease (COPD) is characterized by chronic airway inflammation, associated with extracellular matrix (ECM) degradation and loss of elastic recoil of lung tissue. The *Matrix Metalloprotease* (*MMP*) gene family is thought to participate in the excessive collagenolytic and elastolytic activity that contributes to ECM destruction. MMPs are a family of secreted and membrane associated zinc-dependent endopeptidases, capable of cleaving ECM and non-matrix proteins. Many studies have shown that MMP1, MMP2, MMP9, MMP12 protein and mRNA levels are higher in lung tissue and induced sputum of COPD patients than of controls [1-6]. Proteolytic activities of the MMPs are normally tightly controlled in several ways, e.g. by transcriptional regulation, activation of latent zymogen and interaction with endogenous inhibitors of MMPs, the Tissue Inhibitors of MMPs (TIMPs). Especially the imbalance between MMPs and TIMPs has been proposed to play a major role in ECM destruction and development of COPD, a pulmonary disease strongly associated with smoking. Since virtually all COPD patients have smoked, yet only a subset of smokers develops COPD, it is plausible that the susceptibility to smoking is genetically determined. It is thus reasonable that a genetic origin in the imbalance between MMPs and TIMPs contributes to COPD development.

Single nucleotide polymorphisms (SNPs) have been described in the promoter regions of *MMP1*, *MMP2*, *MMP9* and *MMP12* and they can alter their expression levels [7-10]. Joos *et al.* showed that SNPs in the *MMP1* and *MMP12* promoter regions are more prevalent in subjects with fast FEV_1 decline compared to subjects with no FEV_1 decline in a cohort of current smokers with mild to moderate airway obstruction [11]. The *MMP9* promoter SNP has been associated with emphysema in a case-control study in a Japanese [12], and with COPD in a Chinese population [13]. In contrast, the promoter SNP in *MMP2*, a biologically plausible candidate for COPD, as well as *TIMP1* SNPs have not been studied in relation to COPD development or FEV_1 decline. In *TIMP1*, no SNPs are known that alter its function or expression. Interestingly, two synonymous *TIMP1* SNPs in the gene region responsible for binding and inactivating of MMP9 have been associated with asthma [14]. Given the function of TIMP1, these *TIMP1* SNPs conceivably also play a role in COPD development.

Synthetic MMP inhibitors are being developed as new treatments for COPD, especially since animal studies have shown protective effects of such inhibitors against smoke-induced COPD [15]. Unraveling the genetics of *MMPs* and *TIMPs* in COPD development may identify subjects who may specifically benefit from such treatments. Therefore, we studied SNPs in *MMP1*, *MMP2*, *MMP9*, *MMP12*, and *TIMP1* and their interaction in relation to accelerated FEV_1 decline and COPD development in a general population cohort. To verify our findings, we investigated whether the significant associations could be replicated in an independent cohort of the general population.

METHODS

Subjects

We genotyped DNA from 1390 subjects of the Vlagtwedde/Vlaardingen cohort that participated in the last survey in 1989/1990 [16].This general population-based cohort of Caucasians of Dutch descent started in 1965 and surveys were performed at three year intervals. At each survey, lung function measurements were performed using standardized protocols and questionnaires were completed (see online supplement). The selection of the cohort and details of the study have been described previously [16]. The study protocol was approved by the local university hospital's medical ethics committee and participants gave written informed consent.

As a replication cohort we used data from a random sample of 1152 subjects from the Doetinchem cohort, which is part of the larger MORGEN study [17;18]. The MORGEN study was a random sample of the general population of the Netherlands. Participants of the Doetinchem study underwent spirometry in 1994-1997 and 5 years later in 1999-2003. Characteristics of both study populations are presented in table 1.

Table 1: Characteristics of the Vlagtwedde/Vlaardingen and Doetinchem cohorts

	Vlagtwedde/Vlaardingen N=1390	Doetinchem N=1152
Age, years	52 (35-79)	50 (31-71)
Males, %	51	47
Pack-years	9.0 (0-262.1)	5 (0-84)
last FEV_1%pred*	93.5 (36.0-138.1)	106.6 (39.1-150.7)
ΔFEV_1, ml/yr†	-21.1 (-121;155)	-28.7 (-292; 130)
FEV_1 values, n	7 (1-8)	2 (2-2)
GOLD stage ≥ II, %	13.4	3.2

Data are presented as median (range)

* FEV_1 % predicted at the last surveys of Vlagtwedde/Vlaardingen cohort (1989/1990) and the Doetinchem cohort (1999-2003) † calculated as last-first FEV_1/years participated

Abbreviations: FEV_1 Forced Expiratory Volume in 1 second; GOLD Global initiative for Obstructive Lung Diseases

DNA collection and genotyping

DNA collection and the genotyping protocol of the Vlagtwedde/Vlaardingen study are described in the additional data supplement. We genotyped G-1607GG in *MMP1,* C-1306T in *MMP2,* A-82G and Asn357Ser (A/G) in *MMP12*. For *TIMP1* no tagging SNPs are known, therefore we decided to genotype two noncoding SNPs, previously associated with asthma [14], Phe124Phe (T/C) and Ile158Ile (C/T) in *TIMP1.* With Haploview, using genotype data from the HapMap project [19;20], we selected 3 haplotype tagging SNPs for *MMP9* that tag haplotypes with a frequency above 5% in *MMP9* including 5 kb flanking regions at both the 3'UTR and 5'UTR: rs6065912, rs3918278 and rs8113877. Characteristics of the genotyped SNP are presented in table 2.

The SNPs that were significantly associated with excess FEV_1 decline or COPD development in the Vlagtwedde/Vlaardingen population were genotyped in the Doetinchem cohort by KBioscience (www.kbioscience.co.uk) using a patent-protected system (KASPar). Characteristics of the genotyped SNPs are presented in table 2. We used the statistical software R, "genetics" package (version 1.9.1) to determine whether the SNPs were in Hardy Weinberg equilibrium and linkage disequilibrium (LD).

Table 2: Characteristics of the genotyped SNPs

SNP name	rs number	Chromosome position of gene	Functionality
MMP1 G-1607GG	rs1799750	11q22-q23	G-insertion generates a new 5'-GGA-3' core recognition sequence for members of the E26 transformation-specific (ETS) family of transcription factors
MMP2 C-1306T	rs243865	16q23	T-allele disrupts a Stimulating Protein (SP)-1 binding site, thereby lowering the promoter activity approximately twofold in macrophages and epithelial cells
MMP9 rs6065912	rs6065912	20q11.2-13.1	tagging SNP for *MMP9* gene
MMP9 rs3918278	rs3918278		tagging SNP for *MMP9* gene
MMP9 rs8113877	rs3918278		tagging SNP for *MMP9* gene
MMP12 A-82G	rs2276109	11q22.2-22.5	A-allele has higher affinity for transcription factor Activator protein (AP)-1 and and higher gene expression in reporter gene assays
MMP12 Asn357Ser	rs652438		located in the coding region of the hemopexin domain that is responsible for MMP12 activity, while the function of this polymorphism remains unknown
TIMP1 Phe124Phe	rs4898	Xp11.3-11.23	unknown
TIMP1 Ile158Ile	rs11551797		unknown

Abbreviations: SNP Single Nucleotide Polymorphism; MMP Matrix Metalloprotease; TIMP Tissue Inhibitor of MMP

Statistics

All *TIMP1* analyses were stratified according to sex, since this gene is located on the X-chromosome. To investigate the effect of SNPs on annual FEV_1 decline in the Vlagtwedde/Vlaardingen, we used Linear Mixed Effect (LME) models with adjustment for potential confounders (e.g. sex, first FEV_1 after age 30, pack-years) (see online data supplement) [16;21]. We tested whether there was an interactive effect of *TIMP1* and *MMP* SNPs on FEV_1 decline by introducing their interaction term into the model. We used ANOVA and linear regression models to study SNP effects on first and last available FEV_1 and FEV_1/VC (adjusted for sex, age, pack-years, and height in regression models). We identified subjects with COPD at the last 2 surveys using the GOLD criteria (GOLD stage II

or higher, i.e. $FEV_1/VC< 70\%$ and $FEV_1<80\%$ predicted) [22]. Differences in genotype frequencies of single SNPs for all genes and additionally haplotype frequencies in *MMP9* between subjects with and without COPD were tested using Chi-square tests.

The SNPs that were significantly associated with excess FEV_1 decline or COPD development in the Vlagtwedde/Vlaardingen population were genotyped in the Doetinchem cohort for verification. FEV_1 decline in the Doetinchem cohort was calculated based on FEV_1 decline between the two surveys and genotype effects were tested using linear regression analyses, adjusted for sex, age, pack-years and baseline FEV_1.

Statistical analyses were performed using SPSS (version 14.0.1 for Windows), S-Plus (version 7), the statistical package R (version 1.9.1) [23], and Chaplin [24;25].

RESULTS

Allelic frequencies for the minor alleles of the *MMP* and *TIMP* SNPs in the Vlagtwedde/Vlaardingen population were comparable to those reported in the NCBI dbSNP database: *MMP1* G-1607GG 0.51, *MMP2* C-1306T 0.27, *MMP9* rs3918278 0.03, *MMP9* rs6065912 0.12, *MMP9* rs8113877 0.40, *MMP12* A-82G 0.15, *MMP12* Asn357Ser 0.03, *TIMP1* Ile158Ile in males 0.01, in females 0.01, and *TIMP1* Phe124Phe in males 0.50, and in females 0.53. All SNPs were in Hardy Weinberg equilibrium. The SNPs in *MMP9* were in high LD (pairwise $r^2 >0.8$).

Association of *MMP* SNPs in Vlagtwedde/Vlaardingen

MMP2 C-1306T was significantly associated with accelerated longitudinal decline in FEV_1 in the total population (TT-genotype -4.0 ml/yr excess decline compared to CC-genotype, p=0.027, figure 1), and was also associated with a lower mean FEV_1 % predicted (CC: 92.5, CT: 93.5, TT: 88.5 % predicted; p=0.013) at the last survey. This association remained significant after adjustment for pack-years of smoking in linear regression models. SNPs in *MMP1*, *MMP12* and SNPs and haplotypes in *MMP9* were not associated with longitudinal FEV_1 decline, level of lung function or presence of COPD (GOLD ≥ stage II) (table E1).

Since smoking upregulates MMP activity [26], we also analyzed FEV_1 decline with respect to interaction of the SNPs and smoking. These interaction-terms were not significant.

Association of *TIMP1* SNPs in Vlagtwedde/Vlaardingen

The *TIMP1* Phe124Phe SNP was associated with excess FEV_1 decline in males only (-4.2 ml/yr excess FEV_1 decline compared to wild type (p value 0.041, figure 2). We found that the *TIMP1* Ile158Ile SNP was associated with excess longitudinal FEV_1 decline in both males and females (-30.7 ml/yr respectively -9.5 ml/yr excess FEV_1 decline compared to wild type, p values respectively 0.001 and 0.031, figure 2). The minor allele of the Ile158Ile

SNP was more prevalent in females with COPD than without COPD: CT genotype, 7.3% and 1.5% respectively, p=0.051 (online data supplement, table E1). SNPs in *TIMP1* were not associated with level of lung function cross-sectionally.

Interaction of *TIMP1* and *MMP* on FEV_1 decline in Vlagtwedde/Vlaardingen

We found significant associations of *TIMP1* and *MMP2* SNPs with FEV_1 decline. To test for interaction between these genes, we included interaction terms of *TIMP1* and *MMP2* SNPs in our models on FEV_1 decline, and stratified the analyses by sex. These interaction terms were not significant.

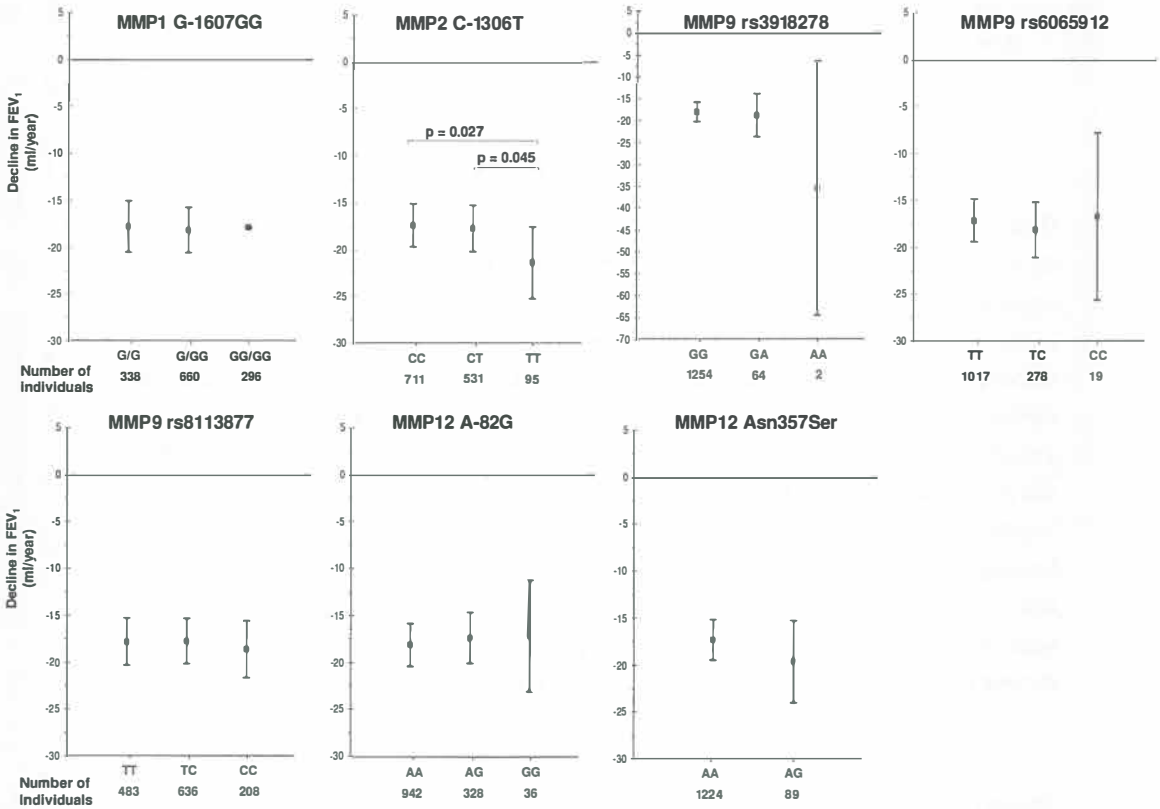


Figure 1: Effect of SNPs in MMP genes on longitudinal decline in FEV_1 *Mean adjusted declines in FEV_1 (in ml/yr) are shown per genotype; bars represent 95% confidence intervals*

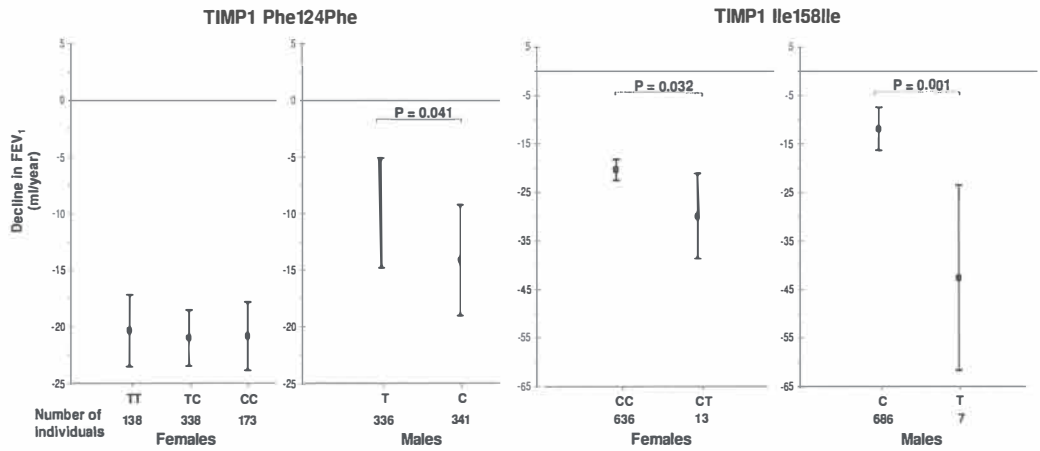


Figure 2: Effect of SNPs in TIMP1 on longitudinal decline in FEV_1, stratified by sex *Mean adjusted declines in FEV_1 (in ml/yr) are shown per genotype; bars represent 95% confidence intervals*

Replication of significant findings in an independent population cohort

To investigate whether results were not found due to chance, we analyzed genes that were significantly associated with FEV_1 level or decline in the Vlagtwedde/Vlaardingen cohort, i.e. *MMP2* and *TIMP1*, in an independent cohort of the general population. Genotype frequencies in the Doetinchem population were similar and not statistically different from the Vlagtwedde/Vlaardingen population (see online supplement, table E2). The *TIMP1* Phe124Phe SNP was associated with excess FEV_1 decline in males (T allele -7.6 ml/yr compared to wild type, p=0.10), similarly to the findings in the Vlagtwedde/Vlaardingen cohort, although with lower significance. In contrast to the findings in Vlagtwedde/Vlaardingen, *TIMP1* Ile158Ile was not associated with excess decline, but with less FEV_1 decline in females (42.9 ml/yr less decline compared to wild type, p=0.008), but not in males. The *MMP2* C-1306T was not significantly associated with excess FEV_1 decline or lower FEV_1% predicted in the Doetinchem cohort.

Discussion

Our study is the first to show that one SNP in *TIMP1* predicts excess FEV_1 decline in two independent populations, implying that the imbalance between MMPs and TIMPs in lung tissue of patients with COPD may be at least partially explained by SNPs that alter the function and/or expression levels of these genes. We could not replicate the association of *MMP2* with excess FEV_1 decline in the second independent population, indicating that the role of genetic variation in *MMP2* on rate of FEV_1 decline is still debatable. In contrast to previous reports from case-control studies, that showed an association of *MMP1*, *MMP9*

and *MMP12* with presence of COPD, decreased levels of FEV_1, and/or excess decline in FEV_1 [12;13], we found no indication whatsoever for a role of *MMP1*, *MMP9* or *MMP12* in the development of COPD or FEV_1 decline in these prospective cohort studies. Consequently, our data suggest that the imbalance in MMPs and TIMPs is likely not disturbed due to genetic variation in the *MMP* genes. This does not rule out that MMPs could play a role in the imbalance in MMPs and TIMPs in COPD. Genetic variation in genes that are involved in regulation of MMPs and TIMPs levels, such as interleukin(IL)-10, IL-13, endothelial growth factor (EGF) and tumor necrosis factor-α (TNF-α) [27;28], may clearly influence the imbalance of MMPs and TIMPs in COPD. Future studies are needed that address the effect of these genes on decline in FEV_1 in the general population.

We show for the first time that genetic variation in *TIMP1* may accelerate the normally occurring FEV_1 decline in the general population. We found that the common SNP Phe124Phe was associated with excess FEV_1 decline in males only. Since the *TIMP1* gene is located on the X-chromosome, carriage of one mutant allele may already account for an effect in males, whereas one mutant allele may be compensated by the other allele in females. Of importance, this association was replicated in the Doetinchem cohort with a larger genotype effect (-9.0 ml/yr excess FEV_1 decline compared to wild type in Doetinchem vs. -4.2 ml/yr excess decline in Vlagtwedde/Vlaardingen), but with somewhat lower significance (p values respectively 0.10 and 0.04 in Doetinchem and Vlagtwedde/Vlaardingen). The Phe124Phe SNP is a synonymous mutation and is therefore unlikely to have a functional effect on protein structure or function. It thus should be regarded as a marker for genetic variation in *TIMP1*. Future studies are warranted to identify SNPs that have a functional effect in this gene. Such SNPs may alter TIMP1 protein structure, resulting in an altered/diminished affinity for MMP9 and subsequently excess MMP9 activity leading to parenchymal destruction.

We observed opposite effects of the *TIMP1* Ile158Ile SNP in the two populations under study. In Vlagtwedde/Vlaardingen the SNP was associated with excess FEV_1 decline in both females and males, whereas in the Doetinchem cohort we found the opposite effect in females (less FEV_1 decline), and no effect was shown in males. The SNP has a very low prevalence and therefore type I errors can easily occur. By testing the SNP in an independent population, we can conclude that the significant effect in the Vlagtwedde/ Vlaardingen population is possibly found by chance.

The *MMP2* C-1306T genotype effect on FEV_1 decline is small in the Vlagtwedde/ Vlaardingen cohort, but we observed no effect at all in the Doetinchem cohort, which may indicate that the association in Vlagtwedde/Vlaardingen may possibly be a spurious result that is not relevant on a population level. On the other hand, we can not completely rule out a genetic effect of *MMP2* since the power to detect small genotype effects is much larger in the Vlagtwedde/Vlaardingen population with a median of 7 survey points compared to 2 survey points in Doetinchem. This may explain the lack of replication.

Further studies with comparable power as in Vlagtwedde/Vlaardingen are warranted to elucidate the role of *MMP2* in FEV_1 decline in the general population.

MMP9 SNPs were not associated with development of COPD or FEV_1 decline in our study. We did not genotype the *MMP9* C-1562T SNP that was previously associated cross-sectionally with the presence of emphysema or COPD in Japanese and Chinese individuals in a case-control study [12;13], due to technical problems. Therefore, we cannot rule out a genetic role of *MMP9* in COPD development. However, we tagged the whole *MMP9* gene for haplotypes with a frequency above 10%, and found strong LD in the whole region. We are therefore confident that we also tagged the C-1562T SNP and that we did not miss information. Alternatively, the causative factor for higher levels of MMP9 in COPD lung tissue does not necessarily have to be genetically determined. *MMP9* is amongst others regulated at the level of transcription by IL-10, IL-13, EGF, and TNF-α [27;28]. It is therefore of interest to analyze SNPs in these genes in the future as well.

We did not confirm previous associations of the *MMP1* and *MMP12* SNPs and lung function decline as previously described by Joos *et al.* [11]. However, in the latter study the *MMP12* Asn357Val SNP was only associated with rate of decline in FEV_1 in combination with the *MMP1* G-1607GG SNP. We performed the same type of analyses and found no association.

To our surprise, we did not find any association of *MMP1, MMP9 and MMP12* SNPs and FEV_1 decline or COPD development in the general population. Although the role of MMPs in COPD pathogenesis has clearly been demonstrated, we are the first to analyze the effects of SNPs in a cohort of the general population, whereas previous studies have used case-control designs. Moreover, differences in phenotypes make the comparison of our study and previous studies difficult. For example, several studies have investigated the effect of the C-1562T SNP in *MMP9* in smokers and nonsmokers on emphysematous phenotypes using chest CT scans [12;13;29]. We do not have CT scans available in Vlagtwedde/Vlaardingen or in Doetinchem, so we can not assess such genetic effects since pulmonary function tests are not very sensitive to detect (mild) emphysema [30].

We feel we did not miss any clinically relevant associations of e.g. 5 ml/year excess FEV_1 decline due to lack of power. For example: for the *MMP9* rs8113877 SNP with approximately 1000 subjects with the wild type genotype and on average a mean annual decline in FEV_1 of 17 ml/yr we have 80% power to detect an excess decline of 5.5 ml/yr in FEV_1 in mutant carriers (n=300), assuming a SD of 3.0 (derived from the actual SE=1.166) in both groups .

In conclusion, our study shows that genetic variation in *TIMP1* is associated with excess FEV_1 decline in two independent general populations. Further research is needed to assess the functionality of this finding. We could not confirm a role for *MMP* SNPs in excess FEV_1 decline and COPD development in the general population, although our study had

sufficient power to detect genetic effects. Since SNPs in *MMP* do no contribute to COPD, it is of interest to assess the genetic contribution of MMP modifying genes, like IL-10, IL-13, EGF, and TNF-α that regulate transcription of *MMPs*. In addition, SNPs in other *TIMPs*, such as *TIMP2*, may also affect the MMP-TIMP balance and thereby exert an effect on FEV_1 decline in the general population.

REFERENCES

1. Demedts IK, Morel-Montero A, Lebecque S, *et al.* Elevated MMP-12 protein levels in induced sputum from patients with COPD. *Thorax* 2006;61(3):196-201.

2. Imai K, Dalal SS, Chen ES, *et al.* Human collagenase (matrix metalloproteinase-1) expression in the lungs of patients with emphysema. *Am J Respir Crit Care Med* 2001;163(3 Pt 1):786-91.

3. Ohnishi K, Takagi M, Kurokawa Y, *et al*. Matrix metalloproteinase-mediated extracellular matrix protein degradation in human pulmonary emphysema. *Lab Invest* 1998;78(9):1077-87.

4. Pons AR, Sauleda J, Noguera A, *et al*. Decreased macrophage release of TGF-{beta} and TIMP-1 in chronic obstructive pulmonary disease. *Eur Respir J* 2005;26(1):60-6.

5. Segura-Valdez L, Pardo A, Gaxiola M, *et al.* Upregulation of gelatinases A and B, collagenases 1 and 2, and increased parenchymal cell death in COPD. *Chest* 2000;117(3):684-94.

6. Boschetto P, Quintavalle S, Zeni E, *et al*. Association between markers of emphysema and more severe chronic obstructive pulmonary disease. *Thorax* 2006;61(12):1037-42.

7. Jorms jo S, Ye S, Moritz J, *et al.* Allele-specific regulation of matrix metalloproteinase-12 gene activity is associated with coronary artery luminal dimensions in diabetic patients with manifest coronary artery disease. *Circ Res* 2000;86(9):998-1003.

8. Price SJ, Greaves DR, Watkins H. Identification of novel, functional genetic variants in the human matrix metalloproteinase-2 gene: role of Sp1 in allele-specific transcriptional regulation. J Biol Chem 2001;276(10):7549-58.

9. Rutter JL, Mitchell TI, Buttice G, *et al.* A single nucleotide polymorphism in the matrix metalloproteinase-1 promoter creates an Ets binding site and augments transcription. *Cancer Res* 1998;58(23):5321-5.

10. Zhang B, Henney A, Eriksson P, *et al.* Genetic variation at the matrix metalloproteinase-9 locus on chromosome 20q12.2-13.1. *Hum Genet* 1999;105(5):418-23.

11. Joos L, He JQ, Shepherdson MB, *et al.* The role of matrix metalloproteinase polymorphisms in the rate of decline in lung function. *Hum Mol Genet* 2002;11(5):569-76.

12. Ito I, Nagai S, Handa T, Muro S, Hirai T, Tsukino M, et al. Matrix metalloproteinase-9 promoter polymorphism associated with upper lung dominant emphysema. *Am J Respir Crit Care Med* 2005;172(11):1378-82.

13. Zhou M, Huang SG, Wan HY, *et al*. Genetic polymorphism in matrix metalloproteinase-9 and the susceptibility to chronic obstructive pulmonary disease in Han population of south China. *Chin Med J* (Engl) 2004;117(10):1481-4.

14. Lose F, Thompson PJ, Duffy D, *et al*. A novel tissue inhibitor of metalloproteinase-1 (TIMP-1) polymorphism associated with asthma in Australian women. *Thorax* 2005;60(8):623-8.

15. Churg A, Wang R, Wang X, *et al.* An MMP-9/-12 Inhibitor Prevents Smoke-induced Emphysema and Small Airway Remodeling in Guinea Pigs. *Thorax* 2007 Aug;62(8):706-13

16. van Diemen CC, Postma DS, Vonk JM, *et al*. A disintegrin and metalloprotease 33 polymorphisms and lung function decline in the general population. *Am J Respir Crit Care Med* 2005;172(3):329-33.

17. Tabak C, Arts IC, Smit HA, *et al.* Chronic obstructive pulmonary disease and intake of catechins, flavonols, and flavones: the MORGEN Study. *Am J Respir Crit Care Med* 2001;164(1):61-4.

18. Grievink L, Smit HA, Ocke MC, *et al.* Dietary intake of antioxidant (pro)-vitamins, respiratory symptoms and pulmonary function: the MORGEN study. *Thorax* 1998;53(3):166-71.

19. The International HapMap Project. *Nature* 2003;426(6968):789-96.

20. Barrett JC, Fry B, Maller J, Daly MJ. Haploview: analysis and visualization of LD and haplotype maps. *Bioinformatics* 2005;21(2):263-5.

21. Pinheiro JC, Bates DM. Mixed-Effects Models in S and S-Plus. New York, NY: Springer; 2000.

22. Fabbri LM, Hurd SS. Global Strategy for the Diagnosis, Management and Prevention of COPD: 2003 update. *Eur Respir J* 2003;22(1):1-2.

23. R development Core Team. R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing; 2004.

24. Case-Control Haplotype Interference (CHAPLIN), version 1.2., 2006.

25. Epstein MP, Satten GA. Inference on haplotype effects in case-control studies using unphased genotype data. *Am J Hum Genet* 2003;73(6):1316-29.

26. Mercer B, Brinckerhoff C, D'Armiento J. Activation of the MMP-1 promoter by cigarette smoke in human small airway epithelial cells requires ERK MAP kinase signaling: differential response of the 1G and 2G promoter sequences. *Proc Am Thorac Soc* 2006 Aug;3(6):477.

27. Lim S, Roche N, Oliver BG, *et al*. Balance of matrix metalloprotease-9 and tissue inhibitor of metalloprotease-1 from alveolar macrophages in cigarette smokers. Regulation by interleukin-10. *Am J Respir Crit Care Med* 2000;162(4 Pt 1):1355-60.

28. Wright JL, Tai H, Wang R, *et al*. Cigarette smoke upregulates pulmonary vascular matrix metalloproteinases via TNF-alpha signaling. *Am J Physiol Lung Cell Mol Physiol* 2007;292(1):L125-L133.

29. Minematsu N, Nakamura H, Tateno H, *et al*. Genetic polymorphism in matrix metalloproteinase-9 and pulmonary emphysema. *Biochem Biophys Res Commun* 2001;289(1):116-9.

30. Gelb AF, Hogg JC, Muller NL, *et al*. Contribution of emphysema and small airways in COPD. *Chest* 1996;109(2):353-9.

Online supplement to paper:

Genetic variation in *TIMP1* but not *MMPs* predict excess FEV_1 decline in the population

Cleo C van Diemen, Dirkje S Postma, Mateusz Siedliński, Henriëtte A Smit, Anneke Blokstra, H Marike Boezen

METHODS

Spirometry

Pulmonary function measurements were performed with a water-sealed spirometer (Lode Spirograph, Lode Instruments, Groningen, The Netherlands). Measurement of inspiratory vital capacity (VC) after a deep expiration was followed by measurement of forced expiratory volume in 1 second (FEV_1). The higher of the values obtained in two technically satisfactory tracings was taken as long as the difference between the two IVC values was less than 150 ml and the difference between two FEV_1 values was less than 100 ml.

DNA collection and Genotyping

During the 1989/1990 survey of the Vlagtwedde/Vlaardingen cohort, neutrophil depot of spinned blood was collected and stored at -20°C. In 2003/2004 DNA was extracted from these samples with the QiaAmp® DNA Blood Mini Kit and checked for purity and concentration with the NanoDrop® ND-1000 UV-Vis Spectrophotometer (Nanodrop Technologies, Wilmington, Delaware, USA).

For all SNPs, primers and probes were obtained from Applied Biosystems TaqMan® SNP Genotyping Assays (Nieuwekerk aan de IJssel, The Netherlands), using the Assay-on demand service. Reactions were performed in 5 µl volumes and contained 10 ng DNA, 1x Taqman Universal Mastermix (Applied Biosystems), 200 nM of each probe and 900 nM of each primer. Cycling conditions on the ABI PRISM® 7900HT Sequence Detection System (Applied Biosystems) were 10 min 95ºC followed by 40 cycles of 15 seconds 95ºC and 1 minute 60ºC. End-point fluorescence was measured immediately after cycling. Alleles were assigned using SDS 2.1 software (Applied Biosystems). Genotyping reactions failed randomly with a maximum percentage of 6% and a minimum of 0%. There was no relationship between degree of missingness and evidence of association. We regenotyped 6% of the samples and found no errors in the genotypes, therefore our data are reliable and reproducible.

Linear Mixed effect models

Time was defined as time in years relative to the first FEV_1, starting from the age of 30. Variables included in the model were sex, pack-years, the first FEV_1 after age 30, and their interaction with time. Since including the level of the first FEV_1 after age 30 and its interaction with time could introduce bias due to regression to the mean, these variables were also included in the model as random effect variables. The results of these analyses showed no change in estimates of the variables in the model or a better fit of the model, which indicates that there was no bias due to regression-to-the-mean. Therefore, the results are presented without these random effects.

Table E1: *MMP* SNPs and COPD development in Vlagtwedde/Vlaardingen (GOLD stage ≥ II)

SNP		No COPD N (%)	COPD N (%)	P value
MMP1 G-1607GG	G	295 (26.4)	44 (24.7)	0.845
	G&GG	565 (50.6)	94 (52.8)	
	GG	257 (23.0)	40 (22.5)	
MMP2 C-1306T	CC	609 (52.8)	103 (55.4)	0.180
	CT	466 (40.5)	65 (34.9)	
	TT	77 (6.7)	18 (9.7)	
MMP12 Asn357Ser	AA	1054 (93.1)	171 (94.0)	0.673
	AG	78 (6.9)	11 (6.0)	
	GG	0 (0)	0 (0)	
MMP12 A-82G	AA	812 (72.2)	130 (71.4)	0.831
	AG	281 (25.0)	48 (26.4)	
	GG	32 (2.8)	4 (2.2)	
MMP9 rs6065912	TT	873 (77.0)	145 (80.1)	0.576
	TC	245 (21.6)	18 (33.2)	
	CC	16 (1.4)	3 (1.7)	
MMP9 rs8113877	TT	408 (35.8)	76 (40.6)	0.137
	TC	559 (49.0)	77 (41.2)	
	CC	174 (15.2)	34 (18.2)	
MMP9 rs3918278	GG	1078 (94.7)	177 (96.7)	0.122
	GA	59 (5.2)	5 (2.7)	
	AA	1 (0.1)	1 (0.6)	

Table E2: *TIMP1* SNPs and development of COPD in Vlagtwedde/Vlaardingen (GOLD stage ≥ II), stratified by sex

		FEMALES					MALES		
		No COPD N (%)	COPD N (%)	P value			No COPD N (%)	COPD N (%)	P value
TIMP1	CC	572 (98.5)	43 (93.5)	0.051	*TIMP1*	C	533 (98.9)	130 (99.2)	0.724
Ile158Ile	CT	9 (1.5)	3 (6.5)		Ile158Ile	T	6 (1.1)	1 (0.8)	
	TT	0 (0)	0 (0)						
TIMP1	TT	122 (21.1)	12 (25.0)	0.298	*TIMP1*	T	264 (50.4)	60 (45.8)	0.348
Phe124Phe	TC	308 (53.2)	20 (41.7)		Phe124Phe	C	260 (49.6)	71 (54.2)	
	CC	149 (25.7)	16 (33.3)						

CHAPTER 7

Decorin and *TGF-β1* polymorphisms and development of COPD in a general population

Cleo C van Diemen
Dirkje S Postma
Judith M Vonk
Marcel Bruinenberg
Ilja M Nolte
H Marike Boezen

Respir Res, 2006;7:89

ABSTRACT

Background

Decorin, an extracellular matrix (ECM) proteoglycan, and TGF-β_1 are both involved in lung ECM turnover. Decorin and TGF-β_1 expression are decreased respectively increased in COPD lung tissue. Interestingly, they act as each other's feedback regulator. We investigated whether single nucleotide polymorphisms (SNPs) in *decorin* and *TGF-β_1* underlie accelerated decline in FEV_1 and development of COPD in the general population.

Methods

We genotyped 1390 subjects from the Vlagtwedde/Vlaardingen cohort. Lung function was measured every 3 years for a period of 25 years. We tested whether five SNPs in *decorin* (3'UTR and four intron SNPs) and three SNPs in *TGF-β_1* (3'UTR rs6957, C-509T rs1800469 and Leu10Pro rs1982073), and their haplotypes, were associated with COPD (last survey GOLD stage ≥ II). Linear mixed effects models were used to analyze genotype associations with FEV_1 decline.

Results

We found a significantly higher prevalence of carriers of the minor allele of the TGF-β1 rs6957 SNP (p=0.001) in subjects with COPD. Additionally, we found a significantly lower prevalence of the haplotype with the major allele of rs6957 and minor alleles for rs1800469 and rs1982073 SNPs in *TGF-β_1* in subjects with COPD (p=0.030), indicating that this association is due to the rs6957 SNP. *TGF-β_1* SNPs were not associated with FEV_1 decline. SNPs in *decorin,* and haplotypes constructed of both *TGF-β_1* and *decorin* SNPs were not associated with development of COPD or with FEV_1 decline.

Conclusions

Our study shows for the first time that SNPs in *decorin* on its own or in interaction with SNPs in *TGF-β_1* do not underlie the disturbed balance in expression between these genes in COPD. *TGF-β_1* SNPs are associated with COPD, yet not with accelerated FEV_1 decline in the general population.

BACKGROUND

Chronic obstructive pulmonary disease (COPD) is characterized by irreversible airway obstruction and persistent airway inflammation. Transforming growth factor-β_1 (TGF-β_1) is one of the important cytokines involved in this inflammatory process, which has been associated with cell proliferation and differentiation. It is furthermore involved in repair of the extracellular matrix (ECM) after inflammation and tissue injury amongst others by promoting synthesis of elastin and collagen. Studies have shown that TGF-β_1 expression is increased in the airways of COPD patients [1;2]. In contrast, a recent article from Pons *et al* showed that alveolar macrophages from COPD patients release less TGF-β_1 in response to lipopolysaccharide than smokers with normal lung function and non-smokers [3]. This may lead to a reduced anti-inflammatory and anti-elastolytic response in COPD patients, subsequently contributing to progressive ECM destruction.

Decorin is a component of the ECM that regulates collagen fibrillogenesis [4-6]. In addition, it can interact with a wide variety of growth factors, cytokines and adhesion molecules through its extensive binding area, thereby not only playing a role in ECM assembly but also in control of cell proliferation and tissue morphogenesis [7]. TGF-β_1 has been shown to downregulate synthesis of decorin in fibroblasts and decorin can in turn inhibit TGF-β_1 [8]. Decorin may thus act as a negative feedback regulator of TGF-β_1 mediated repair responses. Conversely, TGF-β_1 can downregulate expression of decorin in fibroblasts from emphysema patients [9]. We have shown previously that decorin expression is diminished in the peribronchiolar area of lung tissue from patients with severe emphysema, while TGF-β_1 production from fibroblasts of these patients is increased [10]. Noordhoek *et al.* showed that TGF-β_1 and basic fibroblast growth factor give a stronger reduction of decorin production in the culture supernatant of fibroblasts from patients with severe emphysema than from patients with mild emphysema [9]. It thus appears that the regulation of decorin production is disturbed in lung tissue from patients with severe emphysema. This will lead to diminished binding and neutralization of TGF-β_1 by decorin followed by higher TGF-β_1 concentrations and activity with lower decorin production as a result.

We hypothesized that the reciprocal regulation of the *TGF-β_1* and *decorin* genes is disturbed in COPD due to a genetic mutation in one or both of these genes. We have tested this hypothesis by investigating three single nucleotide polymorphisms (SNPs) in *TGF-β_1* and five SNPs in *decorin* on the development of COPD and on lung function decline in a large cohort derived from the general population (the Vlagtwedde/Vlaardingen cohort).

METHODS

Subjects

We used data from 2467 subjects of the Vlagtwedde/Vlaardingen cohort participating in the last survey in 1989/1990. This general population-based cohort of Caucasians of Dutch descent started in 1965. Surveys, during which pulmonary function measurements were performed, were held every three years. The selection of the cohort has been described previously [11;12;13]. Surveys were performed every 3 years during which information was collected on respiratory symptoms, smoking status, age and gender by the Dutch version of the British Medical Council standardized questionnaire. A blood sample was taken and spirometry was performed. Details on pulmonary function measurements are provided in the additional file 1. The methodology for standardization and equipment used for lung function measurements was the same throughout the study. In 1989/1990 neutrophil depot of centrifuged blood was collected and stored at -20°C. In 2003/2004 DNA was extracted from these samples with the QiaAmp® DNA Blood Mini Kit and checked for purity and concentration with the NanoDrop® ND-1000 UV-Vis Spectrophotometer. The study protocol was approved by the local university hospital's medical ethics committee and participants gave written informed consent.

Genotyping

We genotyped DNA of those subjects with more than 1500 ng isolated DNA available (N=1390). Three SNPs, previously associated with COPD or level of lung function were genotyped in *TGF-β_1*: rs6957 in the 3'UTR, rs1800469 in the promoter region (C-509T) and a coding SNP rs1982073 (Leu10Pro, G/T) [14-16]. Coding SNPs in *decorin* have been identified in the NCBI and Celera databases, but are only prevalent in African populations (frequency 0.05-0.12) but not in Caucasian populations (frequency 0.00). According to the HapMap database there are two large LD blocks in the decorin gene, and a region including the 3'UTR that forms no LD block [17]. There are 4 haplotype tagging SNPs located in introns, resulting in 3 major haplotypes, which cover the information of the gene. Therefore, we genotyped one SNP in the 3'UTR (rs1803343), and the 4 haplotype-tagging SNPs: rs11106030, rs741212, rs566806, rs516115 and rs3138241. The genotyping protocol is described in the additional file 1; the characteristics of the genotyped SNPs in additional file 2. To determine whether the SNPs were in Hardy Weinberg equilibrium and whether they were in linkage disequilibrium, tests were performed with the statistical package R (version 1.9.1).

Statistics

We identified subjects with COPD using the GOLD criteria (GOLD stage II or higher, i.e. $FEV_1/VC < 70\%$ and $FEV_1 < 80\%$ predicted) at the last survey [18]. Characteristics of subjects with and without COPD at the last survey are presented in table 1. Differences in allele

Table 1: Characteristics of genotyped subjects in the 1989/1990 survey

	No COPD (N=1156)	COPD (N=186)
Males, n (%)	554 (47.9)	137 (72.9)
Age in years, median (IQR)	50 (35-79)	59 (35-76)
Pack-years of smoking, median (IQR)	7.5 (0-21.6)	25.5 (6.6-35.7)
FEV_1%pred, median (IQR)	95.8 (87.9-104.5)	71.1 (61.1-77.1)
FEV_1/VC, median (IQR)	76.6 (62.1-80.5)	60.0 (54.5-65.7)

Abbreviations: FEV, Forced Expiratory Volume in 1 second; VC Vital Capacity

frequencies and haplotype frequencies between subjects with and without COPD were tested using Chi-square tests. We used ANOVA and linear regression models to study the effect of SNPs on first and last available FEV_1 and FEV_1/VC (adjusted for gender, age, pack-years, and height in regression models).

Linear Mixed Effect (LME) models were used to investigate the effect of SNPs in *TGF-β₁* and *decorin* on annual FEV_1 decline in the general population, like published previously [19;20]. Time was defined as time in years relative to the first FEV_1, starting from the age of 30 [21]. Variables included in the model were age at entry, gender, pack-years, the first FEV_1 after age 30, and their interaction with time. Since including the level of the first FEV_1 after age 30 and its interaction with time could introduce bias due to regression to the mean, these variables were also included in the model as random effect variables. The results of these analyses showed no change in estimates of the variables in the model or a better fit of the model, which indicates that there was no bias due to regression-to-the-mean. Therefore, the results are presented without these random effects. To test whether SNPs were associated with FEV_1 decline within subjects with COPD, we performed LME analyses on these subjects only. Since Celedón *et al.* found stronger linkage results of *TGF-β₁* SNPs and lung function in smokers only, we additionally performed LME models stratified to smoking status [14]. We also included interaction terms of *TGF-β₁* SNPs and *decorin* SNPs to test for genetic interaction of these SNPs.

Instead of performing pre- or post-hoc power analysis and correction for multiple testing, we performed permutation tests to assess whether our results might have been found due to chance. Genotypes were randomly shuffled among individuals to produce 3000 datasets. The LME models were rerun on each of these datasets to generate a distribution of the beta estimates for additional FEV_1 decline of the homozygous minor allele genotype compared to FEV_1 decline of the homozygous wild type allele genotype under the null hypothesis, being no association of the SNPs under study and FEV_1 decline. If the observed beta estimate from the true data is found in the lower 5% percentile of the empiric cumulative distribution ($p<0.05$), one can assume that the observed beta estimate is not found due to chance.

We also estimated *TGF-β1* haplotype frequencies in the whole population and in subjects with a COPD phenotype. Estimated haplotype frequencies for *TGF-β1* higher than 1% in the general population were used to construct phased multi-locus genotypes of *TGF-β1*. For *decorin*, we constructed the phased multi-locus genotypes as known from the HapMap database. With Chi-square tests we determined for each haplotype whether there was a difference in prevalence of carriers between subjects with and without COPD. Also, the excess decline in FEV_1 in the whole population was determined for each phased multi-locus genotype in the LME.

Statistical analyses were performed using SPSS (version 12.0.1 for Windows), the statistical package R (version 1.9.1) and Arlequin [22].

RESULTS

Allelic frequencies for the minor alleles of the *TGF-β1* and *decorin* SNPs in this population were comparable to those reported in the Celera and/or in the NCBI dbSNP database: *TGF-β1* rs6957 0.18, rs1800469 0.28, rs1982073 0.38, *decorin* rs1803343 0.02, rs11106030 0.06, rs741212 0.12, rs566806 0.26, rs516115 0.22 and rs3138241 0.06. All SNPs were in Hardy Weinberg equilibrium. The *TGF-β1* rs1800469 SNP was in significant LD with rs1982073 and rs6957. Rs6957 was in almost significant LD with rs1982072 (p=0.06). The *decorin* SNPs were in significant LD. Graphs of the LD patterns with D', r and P-values in both genes are presented in the additional file 3.

Prevalence of SNPs and haplotypes in *TGF-β1* and *decorin* in COPD and control subjects

The distribution of the *TGF-β1* rs6957 genotypes was significantly different between subjects with and without COPD (p=0.001, table 2). The other *TGF-β1* SNPs were not associated with COPD. We also found no association of SNPs in *decorin* with the prevalence of COPD.

We used estimated haplotype frequencies higher than 0.01 to construct phased multi-locus genotypes for *TGF-β1*. The haplotype consisting of the minor allele for *TGF-β1* rs6957 and the wild type alleles for *TGF-β1* rs1800469 and rs1982073 was more prevalent in subjects with COPD (p=0.014). Because the prevalence of carriers of other haplotypes containing the minor allele at *TGF-β1* rs6957 was also increased in subjects with COPD, this finding only reflects the individual association of the *TGF-β1* rs6957 SNP with COPD. Carriers of at least one haplotype with the minor alleles for *TGF-β1* rs1800469 and rs1982073 and the wild-type allele for rs6957 were less prevalent in COPD (p=0.030). We found no significant associations of phased multi-locus genotypes in *decorin* with the

prevalence of COPD (table 3). We also did not find associations of haplotypes containing SNPs of both *TGF-β1* and *decorin* with COPD (data not shown).

Table 2: Prevalence of genotypes according to COPD phenotype (GOLD stage II or higher; $FEV_1/VC<70\%$, FEV_1 <80% predicted)

SNP		No COPD N (%)	COPD N (%)	P value df=2	SNP		No COPD N (%)	COPD N (%)	P value df=2
TGF-β1	GG	584 (52)	106 (58)	0.541	*Decorin*	AA	878 (76)	131 (76)	0.913
rs1800469	GA	474 (40)	67 (36)		rs741212	AG	242 (22)	43 (22)	
	AA	87 (8)	10 (6)			GG	15 (2)	4 (2)	
TGF-β1	AA	382 (36)	75 (44)	0.297	*Decorin*	AA	614 (55)	102 (55)	0.949
rs1982073	AG	533 (49)	72 (42)		rs516115	AG	431 (38)	65 (38)	
	GG	156 (15)	23 (14)			GG	79 (7)	15 (7)	
TGF-β1	GG	771 (69)	103 (56)	**0.001**	*Decorin*	GG	863 (88)	136 (89)	0.733
rs6957	GA	327 (29)	71 (39)		rs3138241	GA	114 (12)	10 (11)	
	AA	30 (2)	10 (5)			AA	3 (0)	1 (1)	
Decorin	CC	996 (87)	170 (91)	0.217	*Decorin*	AA	1079 (94)	173 (93)	0.507
rs11106030	CA	142 (12)	8 (8)		rs1803343	AG	69 (6)	13 (7)	
	AA	4 (1)	1 (1)			GG	0 (0)	0 (0)	

Abbreviations: COPD Chronic Obstructive Pulmonary Disease; FEV_1 Forced Expiratory Volume in 1 second; VC Vital Capacity; TGF-β1 Transforming Growth Factor-β1; df, degrees of freedom

Table 3: Prevalence of *TGF-β1* and *decorin* haplotypes in subjects with and without COPD (GOLD stage II or higher; $FEV_1/VC<70\%$, FEV_1 <80% predicted)

TGF-β1 : Carrier of Haplotype*			No COPD N (%)	COPD N (%)	P value#
rs1800469	rs1982073	rs6957			
0	0	0	239 (23)	34 (22)	0.686
0	1	0	106 (11)	11 (7.6)	0.264
0	1	1	27 (3)	6 (4)	0.417
1	1	0	288 (29)	31 (20)	**0.030**
0	0	1	95 (9)	25 (16)	**0.014**
1	1	1	160 (16)	34 (22)	0.086
Decorin : Carrier of Haplotype*					
rs3138241	rs516115	rs714212	rs11106030		
0	0	0	0	1009 (93)	175 (92)	0.515
0	1	1	0	234 (22)	47 (27)	0.715
1	1	0	1	133 (12)	15 (9)	0.950

* *0 means wild-type; 1 means minor allele*

\# *P value of Chi-square test for difference in prevalence of haplotype between subjects with and without COPD*

Abbreviations: COPD Chronic Obstructive Pulmonary Disease; FEV_1 Forced Expiratory Volume in 1 second; VC Vital Capacity; TGF-β1 Transforming Growth Factor-β1

Lung function

We found no significant associations (i.e. cross-sectional) between the SNPs tested and FEV_1 and FEV_1/VC at the first or at the last survey in linear regression models (data not shown). The mean adjusted annual decline in lung function (expressed as decrease in FEV_1 in ml/yr) was determined for subjects with the wild-type genotype for the SNPs in *TGF-β1* and *decorin* using LME models. The outcome of the mean annual decline concerns females with age 30 when entered in the LME, a mean first FEV_1 of the population, and zero pack-years. The mean of these adjusted annual declines was 19.2 ml/yr (range 18.7-19.6). We did not find any significant association of SNPs in either *TGF-β1* or *decorin* with accelerated lung function decline (table 4). We added interaction terms of *TGF-β1* en *decorin* SNPs in the model, but found no significant interactions. In addition, we did not find any significant association of haplotypes of either *TGF-β1* or *decorin* with accelerated lung function decline (results not shown). We also tested whether SNPs were associated with lung function decline within subjects with COPD or within smokers, but found no significant associations (table 4 and additional file 4). To test whether results were not missed due to chance, we performed permutation tests. We ran 3000 permutations on our sample of 1390 subjects and performed LME analyses on each of these permutations. The lack of associations with lung function decline was confirmed in these analyses.

Discussion

Decorin and TGF-β_1 can act as each other's feedback regulators in ECM turnover and their expression is respectively decreased and increased in lung tissue of COPD patients. We assessed whether polymorphisms in *decorin* and *TGF-β1* are associated with the development of COPD and accelerated lung function decline in the general population. This is the first study assessing SNPs in *decorin* and we did not find any association with COPD or lung function loss. Contrary to our hypothesis, the observed disturbed balance between decorin and TGB-β_1 in COPD is not caused by a combination of SNPs in their genes, since we found no significant interaction terms of *decorin* and *TGF-β1* SNPs with respect to FEV_1 decline. Moreover, we found no associations of phased multi-locus genotypes containing SNPs of both *TGF-β1* and *decorin* with the presence of GOLD stage II and III COPD in our population. This disturbed balance may be affected by SNPs in *TGF-β1* alone since the 3'UTR SNP in *TGF-β1* is predictive of COPD (stage GOLD II). We found, however, no association of SNPs in *TGF-β1* with longitudinal decline in lung function. In addition, no associations were observed of SNPs in *TGF-β1* with level of FEV_1 or FEV_1/VC cross-sectionally.

It is puzzling that we observed that the *TGF-β1* rs6957 SNP and a haplotype in *TGF-β1* were associated with COPD, but not with excess decline in FEV_1 or with level of FEV_1 and FEV_1/VC at the last survey. We have tested whether there were differences in first

Table 4: Annual decline in FEV_1 according to genotypes of *TGF-β1* and *decorin*

Changes in decline between genotypes in the total population and in subjects who developed COPD (GOLD stage II or higher; $FEV_1/VC<70\%$, FEV_1 <80% predicted) are presented.

		Total population				COPD			
Genotype *TGF-β1*		N	Decline in FEV_1 (ml/yr)*	ΔFEV1 comp to WT	P value†	N	Decline in FEV_1 (ml/yr)*	ΔFEV1 comp to WT	P value†
rs6957	AA	918	-19.2			103	-37.1		
	AG	399	-18.3	+0.9	0.511	71	-33.5	+3.6	0.297
	GG	40	-18.2	+1.0	0.778	10	-28.8	+8.3	0.239
rs1800469	GG	716	-18.9			106	-34.3		
	GA	555	-17.6	+1.2	0.501	67	-36.2	-1.9	0.587
	AA	103	-20.3	-1.5	0.437	10	-31.9	+2.4	0.698
rs1982073	GG	477	-19.1			75	-34.8		
	GA	623	-17.9	+1.2	0.309	72	-33.9	+0.9	0.876
	AA	185	-17.9	+1.2	0.593	23	-35.1	-0.3	0.959
Decorin									
rs1803343	GG	1293	-18.7			173	-35.9		
	GA	85	-18.3	+0.4	0.874	13	-33.6	+2.3	0.698
rs11106030	CC	1206	-18.9			170	-35.2		
	CA	162	-19.6	-0.7	0.688	8	-38.3	-3.1	0.577
	AA	6	-30.5	-11.6	0.285	1	-39.9	-4.7	0.797
rs741212	AA	1039	-18.6			131	-35.1		
	AG	198	-20.1	-1.5	0.287	43	-38.2	-3.1	0.439
	GG	20	-14.1	+4.5	0.346	4	-23.2	+11.9	0.282
rs516115	AA	737	-18.8			102	-34.4		
	AG	519	-18.5	+0.3	0.814	65	-35.9	-1.5	0.669
	GG	96	-18.9	+0.1	0.969	15	-35.0	-0.6	0.930
rs3138241	GG	1187	-18.8			136	-35.7		
	GA	157	-19.5	-0.7	0.694	10	-38.7	-3.0	0.588
	AA	5	-25.7	-6.8	0.589	1	-31.6	+4.1	0.888

**decline in FEV_1 adjusted for gender, first FEV_1 after age 30 years, pack-years, and age; † P value indicates significance of the effect of the genotype on decline in FEV_1 compared to wild-type*

Abbreviations: COPD Chronic Obstructive Pulmonary Disease; FEV_1 Forced Expiratory Volume in 1 second; VC Vital Capacity; TGF-β1 Transforming Growth Factor-β1; WT Wild Type

available FEV_1 (which might suggest a relation to maximal attained lung function level) between the genotypes that could explain the lack of association with FEV_1 decline but this was not the case. Another possibility would be that the FEV_1 decline is only affected by SNPs in certain subgroups, such as smokers. Our stratified analyses showed no such effect.

Although the functionality of the *TGF-β1* rs6957 SNP is not known yet, it has previously been associated with lower pre- and post-bronchodilator FEV_1 and with lower FEV_1/FVC [14]. Similarly, we have shown here that this SNP is associated with development of COPD. Various studies have indicated that the rs1800469 and rs1982073 SNPs are functional and result in higher levels of circulating TGF-β_1 [23-26]. Since TGF-β_1 has anti-inflammatory and

pro-repair activities, these SNPs are thought to be protective against the development of COPD. Indeed, we and others have found that (carriers of haplotypes of) the minor alleles of these SNPs are significantly less prevalent in COPD patients compared to controls [14;16]. Similar to Celedón *et al*, we found an association of a haplotype with at least one minor allele of the rs1800469 and rs1982073 *TGF-β_1* SNPs and COPD, while they also found associations with these SNPs separately [14;16]. The differences in study populations may explain these dissimilarities, e.g. our subjects had milder COPD (FEV_1<80% predicted) than the COPD patients in the Celedón study (FEV_1<45% predicted). Despite the differences in associations, it is still conceivable that carrying both of the SNPs decreases the risk to develop COPD. The two other studies linking *TGF-β_1* SNPs and COPD have also demonstrated that these SNPs are less prevalent in COPD, though these studies did not test haplotypes [15;16].

Many SNPs have been described in the *TGF-β_1* gene, but only a few have been intensively studied in genetic association studies. Cross-sectional studies have found associations of SNPs in *TGF-β_1* with the presence of COPD, and with lower levels of FEV_1 and FEV_1/FVC in several populations [14-16]. We did not analyze every SNP in the *TGF-β_1* gene that was previously reported to be associated with COPD. However, since Celedón *et al*. found strong LD (r^2=0.98) between promoter SNPs and 3'UTR SNPs in a Caucasian population, we are confident that any association that might exist would have been revealed by the SNPs or by their haplotypes [14].

This is the first study on SNPs in *decorin* in a general population or in COPD patients. We were interested in polymorphisms in this gene, since decorin expression in COPD patients is diminished [9;10]. Decorin plays a direct role in the repair processes after inflammation through its regulation of matrix metalloproteases and tissue inhibitors of metalloproteases [27;28]. Furthermore, decorin is the natural inhibitor of TGF-β_1 and may therefore influence the repair process in the lung indirectly. We hypothesized that these processes may be genetically influenced. Since the coding SNPs in decorin described in the NCBI and Celera databases were not prevalent in Caucasians (but only in African populations), we genotyped four tagging SNPs, located in introns, and additionally a 3'UTR SNP. Although we found no significant associations of these SNPs with COPD or lung function decline, we can not rule out completely that there is no genetic defect in *decorin* that increases the risk to develop COPD. However, since we selected tagging SNPs that cover the genetic information of the decorin gene according to HapMap and given the large population under study, we assume that we would have observed an association of SNPs or haplotypes in *decorin* if there existed one in this population.

The lack of a genetic association of SNPs in the *decorin* gene does not rule out an important role of the decorin protein in COPD development. Decorin is a member of the proteoglycan family, a family of macromolecules composed of a protein core with glycosaminoglycan side chains which are produced post-translationally. It is possible that

the function or activation of decorin is disrupted through an altered posttranslational modification of this glycosaminoglycan chain. In this case, modifications in the protein core, which might be caused by SNPs, may not be important and will not be detected. Decorin can be expressed in six splice variants, but the function of these splice variants is not known yet. Nevertheless, a shift in prevalence of one of these splice variants may affect the biological role that decorin exerts in TGF-β_1 regulation, thereby influencing the pathology within the lung.

Conclusions

Contrary to our hypothesis, we were not able to identify the *decorin* gene as a genetic risk factor for the development of COPD. Consequently, SNPs in *decorin* do not seem to underlie a disturbed regulation of this gene and *TGF-β_1* resulting in COPD, nor can they be held responsible for the development of COPD and decline in FEV_1 in the general population. We found that *TGF-β_1* SNPs are associated with the development of COPD but not with accelerated lung function decline or other lung function measures in the general population. Together with previous findings, this study establishes the *TGF-β_1* gene as a risk factor for the development of COPD.

Acknowledgements

This study was funded by the Netherlands Asthma Foundation, grant 3.2.02.51.

REFERENCES

1. Kokturk N, Tatlicioglu T, Memis L, *et al.* Expression of transforming growth factor beta1 in bronchial biopsies in asthma and COPD. *J Asthma* 2003; 40: 887-893.
2. de Boer WI, van Schadewijk A, Sont JK, *et al.* Transforming growth factor beta1 and recruitment of macrophages and mast cells in airways in chronic obstructive pulmonary disease. *Am J Respir Crit Care Med* 1998; 158: 1951-1957.
3. Pons AR, Sauleda J, Noguera A, *et al.* Decreased macrophage release of TGF-beta and TIMP-1 in chronic obstructive pulmonary disease. *Eur Respir J* 2005; 26: 60-66.
4. Svensson L, Heinegard D, Oldberg A. Decorin-binding sites for collagen type I are mainly located in leucine-rich repeats 4-5. *J Biol Chem* 1995; 270: 20712-20716.
5. Schonherr E, Hausser H, Beavan L, *et al.* Decorin-type I collagen interaction. Presence of separate core protein-binding domains. *J Biol Chem* 1995; 270: 8877-8883.
6. Bhide VM, Laschinger CA, Arora PD, *et al*. Collagen phagocytosis by fibroblasts is regulated by decorin. *J Biol Chem* 2005; 280: 23103-23113.
7. Iozzo RV, Murdoch AD. Proteoglycans of the extracellular environment: clues from the gene and protein side offer novel perspectives in molecular diversity and function. *FASEB J.* 1996; 10: 598-614.
8. Yamaguchi Y, Mann DM, Ruoslahti E. Negative regulation of transforming growth factor-beta by the proteoglycan decorin. *Nature* 1990; 346: 281-284.
9. Noordhoek JA, Postma DS, Chong LL, *et al.* Different modulation of decorin production by lung fibroblasts from patients with mild and severe emphysema. *COPD: Journal of Chronic Obstructive Pulmonary Disease* 2005; 2: 17-25.
10. van Straaten JF, Coers W, Noordhoek JA, *et al*. Proteoglycan changes in the extracellular matrix of lung tissue from patients with pulmonary emphysema. *Mod Pathol* 1999; 12: 697-705.
11. Rijcken B, Schouten JP, Mensinga TT, *et al.* Factors associated with bronchial responsiveness to histamine in a population sample of adults. *Am Rev Respir Dis* 1993; 147: 1447-1453.
12. Van der Lende R, Kok T, Peset R, *et al.* Longterm exposure to air pollution and decline in VC and FEV1. Recent results from a longitudinal epidemiologic study in the Netherlands. *Chest* 1981; 80: 23-26.
13. Van der Lende R, Orie NG. The MRC-ECCS questionnaire on respiratory symptoms (use in epidemiology). *Scand J Respir Dis* 1972; 53: 218-226.
14. Celedon JC, Lange C, Raby BA, *et al*. The transforming growth factor-beta1 (TGFB1) gene is associated with chronic obstructive pulmonary disease (COPD). *Hum Mol Genet* 2004; 13: 1649-1656.
15. Su ZG, Wen FQ, Feng YL, *et al.* Transforming growth factor-beta1 gene polymorphisms associated with chronic obstructive pulmonary disease in Chinese population. *Acta Pharmacol Sin* 2005; 26: 714-720.
16. Wu L, Chau J, Young RP, Pokorny V, *et al.* Transforming growth factor-beta1 genotype and susceptibility to chronic obstructive pulmonary disease. *Thorax* 2004; 59: 126-129.
17. The International HapMap Consortium: The International HapMap Project. *Nature* 2003; 426: 789-796.
18. Fabbri LM, Hurd SS. Global Strategy for the Diagnosis, Management and Prevention of COPD: 2003 update. *Eur Respir J* 2003; 22: 1-2.

19. van Diemen CC, Postma DS, Vonk JM, *et al*. A disintegrin and metalloprotease 33 polymorphisms and lung function decline in the general population. *Am J Respir Crit Care Med* 2005; 172: 329-333.

20. Pinheiro JC, Bates DM: *Mixed-Effects Models in S and S-Plus*.Springer, New York, NY, 2000.

21. Rijcken B, Weiss ST. Longitudinal analyses of airway responsiveness and pulmonary function decline. *Am J Respir Crit Care Med* 1996; 154: S246-S249.

22. R development Core Team, *R: A language and environment for statistical computing*.R Foundation for Statistical Computing, Vienna, Austria, 2004.

23. Silverman ES, Palmer LJ, Subramaniam V, *et al*. Transforming Growth Factor-β1 Promoter Polymorphism C-509T Is Associated with Asthma. *Am J Resp Crit Care Med.* 2004; 169: 214-219.

24. Hobbs K, Negri J, Klinnert M, *et al.* Interleukin-10 and transforming growth factor-beta promoter polymorphisms in allergies and asthma. *Am J Respir Crit Care Med* 1998; 158: 1958-1962.

25. Grainger DJ, Heathcote K, Chiano M, *et al.* Genetic control of the circulating concentration of transforming growth factor type beta1. *Hum Mol Genet* 1999; 8: 93-97.

26. Awad MR, El Gamel A, Hasleton P, *et al.* Genotypic variation in the transforming growth factor-beta1 gene: association with transforming growth factor-beta1 production, fibrotic lung disease, and graft fibrosis after lung transplantation. *Transplantation* 1998; 66: 1014-1020.

27. Imai K, Hiramatsu A, Fukushima D, *et al*. Degradation of decorin by matrix metalloproteinases: identification of the cleavage sites, kinetic analyses and transforming growth factor-beta1 release. *Biochem J* 1997; 322 (Pt 3): 809-814.

28. Haj Zen A, Lafont A, Durand E, *et al.* Effect of adenovirus-mediated overexpression of decorin on metalloproteinases, tissue inhibitors of metalloproteinases and cytokines secretion by human gingival fibroblasts. *Matrix Biol* 2003; 22: 251-258.

Additional files to paper:

Decorin and *TGF-β*, polymorphisms and development of COPD in a general population

Cleo C van Diemen, Dirkje S Postma, Judith M Vonk, Marcel Bruinenberg, Ilja M Nolte, H Marike Boezen

ADDITIONAL FILE 1 – METHODS

Pulmonary function measurements

Pulmonary function measurements were performed with a water sealed spriometer (Lode Spirograph D53, Lode Instruments, Groningen, The Netherlands). An inspiratory vital capacity (IVC) was measured after a deep expiration and followed by measurement of forced expiratory volume in 1 second (FEV_1). Subjects performed the manoeuvre until two technically satisfactory tracings were produced. The higher value of the two tracings was taken as the baseline measurement. For a tracing to be acceptable, the difference between two IVC values could not be more than 150 ml, and the difference between two FEV_1 measurements not more than 100 ml. All values were recorded at ATPS. The surveys always took place in October.

Genotyping protocol

Primers and probes of all SNPs were obtained from Applied Biosystems TaqMan® SNP Genotyping Assays (Nieuwekerk aan de IJssel, The Netherlands), using the Assay-by-Design service, for which we provided sequences, or the Assay-on-demand service when assays were already designed by Applied Biosystems. Reactions were performed in 5 µl volumes and contained 10 ng DNA, 1x Taqman Universal Mastermix (Applied Biosystems), 200 nM of each probe and 900 nM of each primer. Cycling conditions on the ABI PRISM® 7900HT Sequence Detection System (Applied Biosystems) were 10 min 95°C followed by 40 cycles of 15 seconds 95°C and 1 minute 60°C. End-point fluorescence was measured after cycling. Alleles were assigned using SDS 2.1 software (Applied Biosystems). Genotyping reactions failed randomly with a minimum of 0.02% and a maximum percentage of 6.4% per SNP. There was no relationship between degree of missing genotypes and evidence of association. We regenotyped 6% of the samples and found no errors in any of the genotypes.

ADDITIONAL FILE 2 – CHARACTERISTICS OF GENOTYPED SNPS

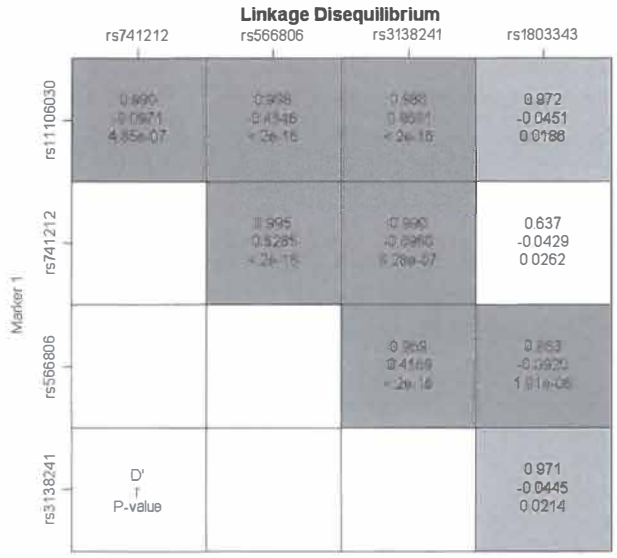
SNP name	rs number	Location of SNP in gene	MAF	Functionality	Primers & probes*
TGF-β1 C-509T	rs1800469	promoter -509	0.282	associated with increased levels of TGF-β_1	For: GGAGAGCAATTCTTACAGGTGTCT Rev: GGAGAAGAGGGTCTGTCAACATG Vic: ACACCTGAGGGATGG Fam: ACACCTGAAGGATGG
TGF-β1 Leu10Pro	rs1982073	exon 1, +29	0.376	associated with increased levels of TGF-β_1 in serum	For: CGCGCTCTCGGCAGT Rev: AGGCGTCAGCACCAGTAG Vic: CAGCAGCGGCAGCA Fam: CAGCAGCAGCAGCA
TGF-β1 3'UTR	rs6957	locus 28344 3'UTR	0.178	unknown	AB assay C_7818385_10
Dec_int1	rs741212	intron before exon1	0.120	unknown	AB assay C_2309576_10
Dec_int2-3	rs566806	intron between exons 2-3	0.262	unknown	AB assay C_2675824_10
Dec_int3-4	rs516115	intron between exons 3-4	0.223	unknown	For: GGGCATGGGAACCAACAAG Rev: GCTCTGAGACCCTCAAATTCTTCTA Vic: CACTGTGAAGCACCCA Fam: CACTGTGAGGCACCCA
Dec3'UTR	rs1803343	3'UTR	0.024	unknown	AB assay C_7561184_10

**Sequences of primers and probes are listed when the assay was designed by Applied Biosystems (AB) assay by design, otherwise sequences are known by AB*

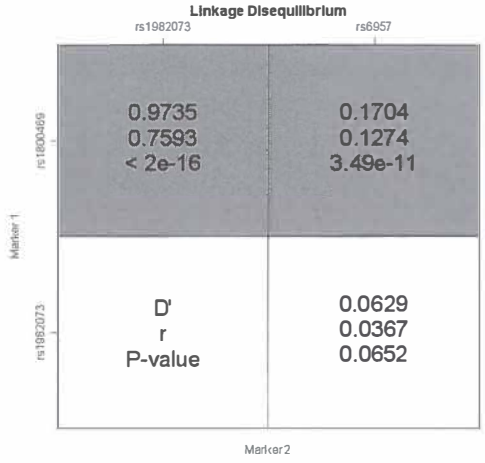
Abbreviations: SNP, Single Nucleotide Polymorphism; MAF, Minor Allele Frequency; AB assay, Applied Biosystems assay on demand; TGF-β_1, Transforming Growth Factor-β_1; UTR, Untranslated Region

ADDITIONAL FILE 3: LINKAGE DISEQUILIBRIUM OF SNPS IN *DECORIN* AND *TGF-B_1*

Decorin:



TGF-B_1:



ADDITIONAL FILE 4: ASSOCIATION OF ANNUAL DECLINE IN FEV_1 ACCORDING TO GENOTYPES OF *TGF-B_1* AND *DECORIN*

Changes in decline between genotypes in never smokers and current and past smokers.

Genotype *TGF-B_1*		Decline in FEV_1 (ml/yr)* Never smokers	ΔFEV_1 compared to WT	P value†	Decline in FEV_1 (ml/yr)* Ever smokers	ΔFEV_1 compared to WT	P value†
rs6957	AA	-20.4			-21.7		
	AG	-21.0	-0.6	0.736	-19.5	+2.2	0.230
	GG	-25.3	-4.9	0.312	-18.1	+3.6	0.426
rs1800469	GG	-21.3			-20.8		
	GA	-20.2	+1.1	0.512	-19.6	+1.2	0.498
	AA	-20.6	+0.7	0.831	-19.6	+1.2	0.712
rs1982073	GG	-21.3			-20.2		
	GA	-20.0	+1.3	0.588	-19.2	+1.0	0.673
	AA	-20.9	+0.4	0.864	-22.4	-2.2	0.402
Decorin							
rs1803343	GG	-21.0			-30.3		
	GA	-23.2	-2.2	0.469	-17.7	+2.6	0.421
rs11106030	CC	-20.9			-20.4		
	CA	-21.2	-0.3	0.918	-21.1	-0.7	0.793
	AA	-29.5	-8.6	0.594	-29.8	-9.3	0.511
rs741212	AA	-21.2			-19.9		
	AG	-20.5	+0.7	0.708	-22.7	-2.8	0.165
	GG	-18.3	+2.9	0.533	-9.5	+10.4	0.246
rs566806	AA	-21.3			-20.9		
	AG	-21.3	-0.0	0.992	-20.1	+0.8	0.639
	GG	-20.2	+1.1	0.691	-19.6	+1.3	0.701
rs516115	AA	-21.4			-20.5		
	AG	-21.1	+0.3	0.873	-19.6	+0.9	0.613
	GG	-20.6	+0.8	0.779	-20.7	-0.2	0.849
rs3138241	GG	-20.9			-20.3		
	GA	-21.2	-0.3	0.881	-20.8	-0.5	0.865
	AA	-29.3	-8.4	0.598	-21.0	-0.7	0.969

** decline in FEV_1 adjusted for gender, first FEV_1 after age 30 years, and age; † P value indicates significance of the effect of the genotype on decline in FEV_1 compared to wild-type*

Abbreviations: COPD Chronic Obstructive Pulmonary Disease; FEV_1 Forced Expiratory Volume in 1 second; VC Vital Capacity; TGF-B_1 Transforming Growth Factor-B1; WT Wild-Type

CHAPTER 8

GSTM1 and *GSTT1* null alleles are risk factors for excess lung function decline: a longitudinal cohort study of the general population

Cleo C van Diemen
H Marike Boezen
Marcel Bruinenberg
Dirkje S Postma

Submitted for publication

ABSTRACT

Background

Cigarette smoke induced oxidative stress plays an important role in the onset of Chronic Obstructive Pulmonary Disease (COPD). Males and females respond differently to cigarette smoke. Polymorphisms of the antioxidative and detoxifying *Glutathione S-Transferase* (*GST*) genes have been associated with lung function decline and COPD.

Aim

To study whether null alleles of *GSTM1* and *GSTT1* and Single Nucleotide Polymorphisms (SNPs) in *GSTP1* are associated with COPD development and excess FEV_1 decline in the general population, and whether associations are affected by gender and smoking.

Methods

1390 Caucasians from a longitudinal cohort study were genotyped for null alleles of *GSTM1* and *GSTT1*, and SNPs Ile105Val and Ala114Val in *GSTP1*. The risk to develop COPD (defined as GOLD stage ≥ II) was analyzed using logistic regression. FEV_1 decline according to genotype was analyzed using linear mixed effect models. All analyses were adjusted for appropriate confounders. We additionally performed stratified analyses according to smoking and sex.

Results

Subjects heterozygous for the *GSTT1* null allele had a significantly higher risk to develop COPD (odds ratio 2.03, 95% CI 1.29 - 3.19, p=0.002) and also a significant excess FEV_1 decline (-17.3 vs. -20.7 ml/yr, p=0.035) compared to subjects with homozygous wild type null allele. Subjects homozygous for the *GSTM1* null allele had excess FEV_1 decline compared to carriers of at least one wild type allele (-16.5 vs. -18.9 ml/yr, p=0.023). Stratified analyses according to smoking (never vs. ever) and sex showed that only ever smoking males with the *GSTT1* null allele had a significant excess FEV_1 decline. This was also the case for ever smoking males with the *GSTM1* null allele. The *GSTM1* and *GSTT1* null alleles were not associated with excess FEV_1 decline in females, irrespective whether we stratified according to smoking habits.

Conclusions

The *GSTM1* and *GSTT1* null alleles confer risk for excess FEV_1 decline. Especially males have an increased risk when carrying these null alleles, an effect that is enhanced by smoking. The *GSTT1* null allele additionally confers risk to develop COPD. *GSTP1* SNPs do not appear to play a role in COPD development or FEV_1 decline in the general population.

INTRODUCTION

Chronic Obstructive Pulmonary Disease (COPD) is characterized by irreversible airway obstruction, chronic airway inflammation, and degradation of lung tissue. The pathogenesis of COPD encompasses a number of injurious processes, including an abnormal inflammatory response in the lungs to inhaled cigarette smoke and other particles and gases. The major toxic compounds in cigarette smoke are the reactive oxygen species, inducing oxidative stress [1-3]. There is solid support for increased lung and systemic oxidative stress in COPD, based on markers in peripheral blood cells, bronchoalveolar lavage (BAL) and sputum [4]. The importance of oxidative stress has recently become more eminent given its link with cellular damage/response and transcriptional control of inflammation [5]. There are several endogenous antioxidant defenses that counterbalance the effects of oxidants, amongst others the glutathione system. Glutathione is the most abundant non-protein thiol in BAL and its levels are increased in smokers [6]. Glutathione can neutralize small reactive oxygen species by itself, but can also be conjugated to larger xenobiotic compounds by enzymes called Glutathione S-Transferases (GST).

There are some well-known functional polymorphisms in *GST* genes, i.e. *GSTM1*, *GSTT1*, *GSTP1*, and especially deletions are of importance. Null alleles have been described for *GSTM1* and *GSTT1*, resulting from a partial gene deletion. The deletions cause an enzymetic defect in detoxification of carcinogens, especially polycyclic aromatic hydrocarbons present in cigarette tar [7;8]. The *GSTM1* null allele has been associated with clinical symptoms of chronic bronchitis in smokers [9], development of COPD [10] and lower lung function growth in children [11]. The *GSTT1* null allele alone or in combination with the *GSTM1* homozygous gene deletion has been associated with excess decline in FEV_1, both in smokers and non smokers in the general population [12]. The single nucleotide polymorphism (SNP) in the *GSTP1* gene, Ile105Val, is known to affect the catalytic efficiency of the GSTP1 enzyme for certain compounds in cigarette smoke. Enzymes with valine at 105 have 7-fold higher catalytic efficiency for diol epoxides of polycyclic aromatic hydrocarbons, but are 3-fold less effective using 1-chloro-2,4,-dinitrobenzene as a substrate [13]. Interestingly, genotypes containing Ile at position 105 are more prevalent in COPD [14]. For the Ala114Val substitu-tion in *GSTP1*, no functional effect has been described, but it has been suggested to aug-ment the Ile105 substitution and it has been associated with lower lung function and COPD [15].

In the current study, we aimed to expand the knowledge on genetic effects of the described polymorphisms in *GSTs* on lung function. We investigated the effect of the null alleles of *GSTM1* and *GSTT1* and the Val105Ile and Ala114Val SNPs in *GSTP1* on longitudinal decline in lung function and development of COPD in a large prospective cohort of the general population. Since the GSTs play a particular role in detoxification of cigarette smoke, we tested for interaction of the genotypes and smoking. Additionally, a previous

study by Imboden *et al.* demonstrated effect modification of the *GSTM1* and *GSTT1* null alleles by sex [12], we therefore performed additional analyses on interaction of *GST* alleles and sex.

METHODS

Subjects

We genotyped DNA from 1390 subjects of the Vlagtwedde/Vlaardingen cohort that participated in the last survey in 1989/1990 [16;17].This general population-based cohort of Caucasians of Dutch descent started in 1965 and surveys were performed at 3-year intervals. At each survey, lung function measurements were performed using standardized protocols and questionnaires were completed. The selection of the cohort and details of the study have been described previously [16;17]. The study protocol was approved by the local university hospital's medical ethics committee and participants gave written informed consent. Characteristics of all subjects are presented in table 1.

Table 1: Characteristics of the study population Median values (interquartile range) are presented

	Total population	Ever smokers	Never smokers	Males	Females
Number of individuals	1390	945	445	714	676
Age, years	52 (43 – 60)	50 (43-59)	53 (44-61)	50 (42-59)	53 (44-60)
Sex, %male	51.4	65.5	23.3	100	0
Smokers					
% Ever	67.8	100	0	85.4	49.6
% Never	32.2	0	100	14.6	50.4
Pack-years smoking	9.0 (0 – 24.0)	18.9 (8.7-30.9)	0	20.4 (6.9-32.9)	0 (0-10.7)
FEV_1%pred	93.5 (83.8-104.0)	91.4 (82.0-100.4)	97.9 (88.6-106.2)	90.6 (80.9-99.7)	96.6 (87.1-105.2)
FEV_1/VC, %	75.1 (69.6-79.8)	73.4 (67.8-78.8)	77.8 (73.7-81.6)	73.1 (66.7-78.3)	77.2 (71.7-80.9)
FEV_1 values, n	7 (5-8)	7 (5-8)	7 (5-8)	7 (5-8)	7 (5-8)

Abbreviations: FEV_1: Forced Expiratory Volume in 1 second;VC: Vital Capacity

DNA collection and genotyping

DNA collection and the genotyping protocol are described in the additional data supplement. We genotyped *GSTP1* Ile105Val (rs1695) and Ala114Val (rs1138272) using TaqMan® SNP Genotyping Assays (Applied Biosystems, USA). For *GSTM1* and *GSTT1* null alleles we designed primers (Biolegio BV, Nijmegen, The Netherlands) and probes (Biolegio BV, Applied Biosystems, Warrington, UK) for multiplex real-time PCR, using β-globulin as internal control. Primers, probes and the genotyping protocols are available upon request. PCR reactions were performed on the Taqman® 7900 HT system (Applied Biosystems).

Statistics

We identified subjects with COPD at the last 2 surveys using the GOLD criteria (GOLD stage II or higher, i.e. $FEV_1/VC< 70\%$ and $FEV_1<80\%$ predicted) [18]. We used logistic regression analyses to test for association of the genotypes with COPD development, adjusted for sex, age, pack-years of smoking. Linear Mixed Effect (LME) models were used to investigate the effect of SNPs on annual FEV_1 decline, with adjustment for potential confounders sex, first FEV_1 after age 30, and pack-years of smoking [16;19]. We tested for interaction of the genotypes and smoking by introducing the interaction term genotype *smoking status (ever vs. never smokers) in the model. Similarly, we tested for interaction between genotypes and sex by introducing the interaction term genotype*sex in the models. When we found a significant interaction, we performed stratified analyses according to smoking status and/or sex to analyze the effects of genotypes within certain groups. We additionally tested whether carrying a combination of *GSTM1* and *GSTT1* null alleles was associated with increased risk for COPD or excess FEV_1 decline. Statistical analyses were performed using SPSS (version 14.0.1 for Windows), S-Plus (version 7) and the statistical package R (version 1.9.1) [20].

Results

The allelic frequencies of the *GST* polymorphisms were: *GSTM1* null 0.51; *GSTT1* null 0.08, *GSTP1* Ile105Val 0.39, *GSTP1* Ala114Val 0.10. All polymorphisms were in Hardy Weinberg equilibrium.

COPD development

The *GSTT1* null allele heterozygous genotype was associated with increased risk to develop COPD compared to wild type, adjusted for age, sex, and pack-years smoking (OR 2.02, 95% CI 1.29-3.19) (table 2). In a dominant model the *GSTT1* null allele was significantly associated with COPD development with an OR 1.65 (95% CI 1.10-2.48, p=0.016) of the heterozygous and double deletion compared to homozygous wild type. *GSTM1* null allele, *GSTP1* Ile105Val, *GSTP1* Ala114Val and the combination of *GSTM1* and *GSTT1* null alleles were not associated with COPD (table 2).

Decline in lung function

The *GSTM1* and *GSTT1* null alleles were significantly associated with excess FEV_1 decline in the total general population compared to the wild type alleles (*GSTM1*: wild type -16.5 vs. null allele -18.9 ml/yr, p=0.023 and *GSTT1*: wild type -17.3 vs. heterozygous genotype -20.7 ml/yr, p=0.035, see figure 1). There existed a significant interaction between the *GSTM1* and *GSTT1* null alleles and smoking (p<0.05).

Table 2: Association of *GST* SNPs and COPD (GOLD stage ≥ II) development in the Vlagtwedde/ Vlaardingen cohort Odds ratios with 95% confidence intervals are presented for the risk to develop COPD compared to wild type, adjusted for age, sex and pack-years

		OR	95% CI	P value
GSTM1 null	0/0	1.03	0.73 - 1.45	0.873
GSTT1 null	wt/0	2.03	1.29 - 3.19	**0.002**
	0/0	0.89	0.39 - 2.04	0.783
GSTP1 Ile105Val	AG	1.12	0.79 - 1.61	0.525
	GG	0.81	0.47 - 1.38	0.435
GSTP1 Ala114Val	CT	0.99	0.64 - 1.54	0.946
	TT	1.67	0.43 - 6.43	0.457

Abbreviations: GST Glutathion S transferase; COPDChronic Obstructive Pulmonary Disease; GOLD Global initiative for Obstructive Lung Diseases; OR Odds Ratio; CI Confidence Interval; wt wild type

Stratified analyses according to ever versus never smoking showed a significant association of the *GSTM1* null allele with excess FEV_1 decline in ever smokers, but not in never smokers (figure 1). For the *GSTT1* null allele, we found a borderline significant effect on excess FEV_1 decline in ever smokers (p=0.098), whereas no effect was shown in never smokers. The combination of *GSTM1* and *GSTT1* null alleles was not associated with excess FEV_1 decline.

When stratifying according to sex, the *GSTM1 and GSTT1* null alleles were associated with accelerated FEV_1 decline in males, but not in females. Stratified analyses in males according to ever versus never smokers showed a significant association of the *GSTM1* null allele in male ever smokers, but not in never smokers (figure 2). We did not find significant effects in females, nor when the analyses were stratified according to smoking.

Discussion

This study shows significant associations of *GSTT1* and *GSTM1* null alleles with excess lung function decline in the general population. Interestingly, we found that these genetic effects were modified by smoking and sex. The *GSTM1* and *GSTT1* null alleles were associated with excess FEV_1 decline in male smokers exclusively and do not appear to affect FEV_1 decline in females at all.

Our study confirms the general message of the recent study by Imboden *et al.*, which demonstrated gender and smoking modification of *GST* null alleles on FEV_1 decline [12]. We can confirm certain conclusions from their studies, but we also demonstrate different results. We showed a significant effect of the *GSTT1* null allele in the general population, but this effect was actually driven by the association in males, in whom significant effects on FEV_1 decline were observed, which were not present in females. The association in males remained significant within male ever smokers, but not in male never smokers, which indicates a modifying effect of smoking on the genetic effect. The results from

Imboden *et al*. demonstrated that the *GSTT1* null allele was associated with FEV_1 decline in males independently of smoking. Our results are thus somewhat different. Notwithstanding this, we did find a larger genetic effect in never smokers than in ever smokers in our population (respectively -6.1 ml/yr vs. -14.6 ml/yr of the heterozygous allele compared to wild type), but this did not reach significance. This may be explained by a lack of power to detect differences between genotypes, since the group of male never smokers was small (n=104).

A major difference between our study and the study by Imboden *et al*. is that we found a significant effect of *GSTM1* null in the total general population, whereas they only showed a significant effect when the *GSTM1* null genotype was combined with the *GSTT1* null genotype. In our study, the effect of the *GSTM1* null allele on FEV_1 decline in the total population is likely caused by the association in male ever smokers who displayed excess FEV_1 decline (-5.0 ml/yr excess decline compared to wild type). A similar effect was shown in female ever smokers (-2.8 ml excess decline), but this was not significant.

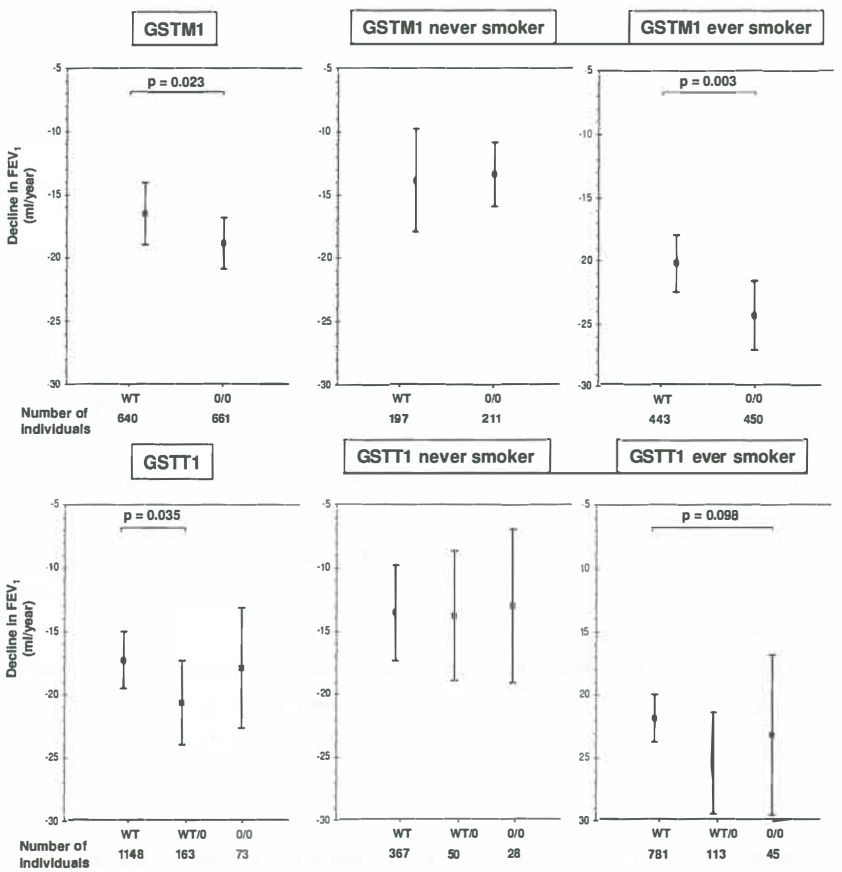


Figure 1: Association of GSTM1 and GSTT1 null alleles with FEV_1 decline in the total Vlagtwedde/Vlaardingen population, and stratified by smoking status. *Mean declines per genotype are presented adjusted for sex, baseline FEV_1 and pack-years of smoking (when appropriate), with 95% CI.*

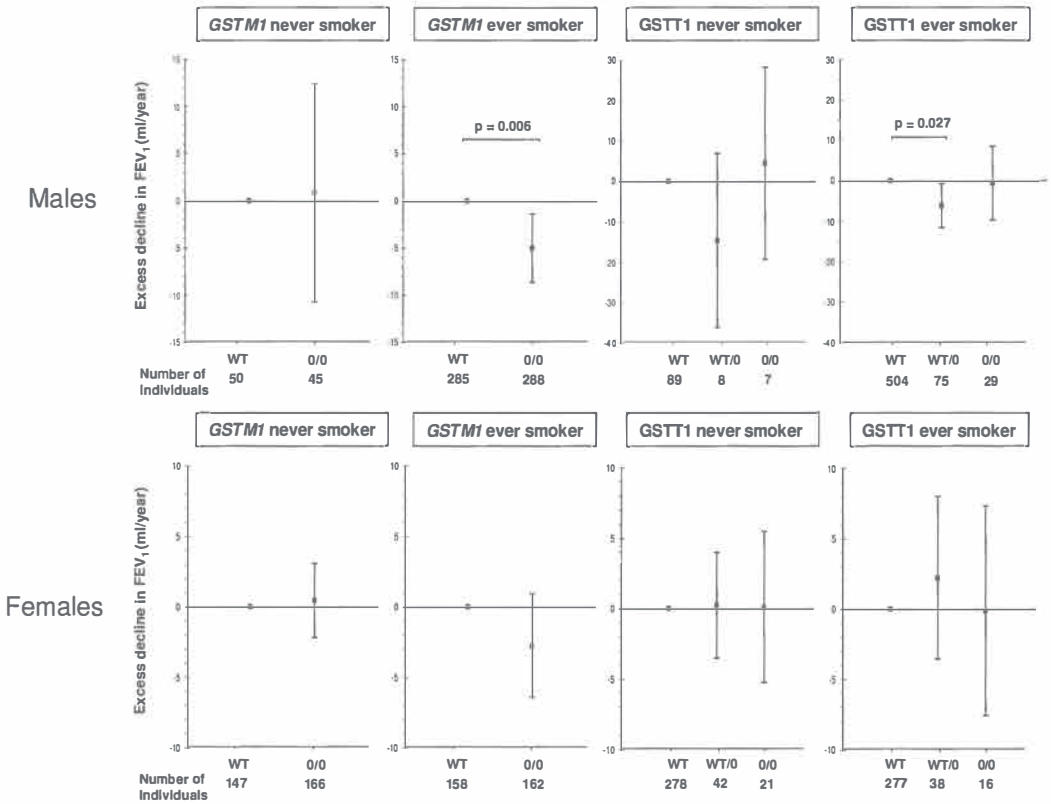


Figure 2: Association of GSTM1 and GSTT1 null alleles with excess FEV_1 decline stratified by sex and smoking status. The wild type (WT) genotype is set at zero, and excess declines per genotype are presented, adjusted for baseline FEV_1 and pack-years of smoking (when appropriate), with 95% CI.

The difference in effects of the *GSTT1* and *GSTM1* null alleles in males compared to females may be partially explained by the fact that more males had smoked and they had smoked more pack-years than females (see table 1). Thus, males in our study were more at risk to smoke induced FEV_1 decline if they missed the protective effects of antioxidant enzyme activity due to the *GSTM1* and *GSTT1* null alleles. Furthermore, it has been suggested that women are better protected against oxidative stress since they display higher baseline levels of antioxidants in plasma compared to men [21]. Another protective mechanism may be that women have been reported to consume more food with high antioxidants content [22]. It is unknown how much of digested antioxidants reach the lung, and to what extent this protects the lung. However, a high vitamin C (a potent anti-oxidant) intake has been associated with a higher lung function in some studies, but not all [22-24]. The better protection against oxidative stress in women may thus compensate the lack of GSTM1 and GSTT1 enzyme activity. We do not have dietary information of our general population; therefore we can not assess such effects. Further studies in larger

populations are warranted to assess these gene-environment interactions in males and females.

Although the null alleles of *GSTM1* and *GSTT1* do not appear to affect FEV_1 decline or risk to develop COPD in women, even not when they smoke, this does not refute the general held opinion that females are more susceptible to develop COPD [25-27]. There were fewer females with COPD in our cohort, yet they developed COPD with lower doses of cigarette smoke compared to males. In addition women may display a different phenotype of COPD compared to men. It has been demonstrated that women have less severe overall emphysema, particularly in the periphery of the lung, but that they have thicker airway walls and smaller lumens, whereas men display more emphysema [28;29]. These results suggest that men and women may respond differently in the type and location of lung damage due to tobacco exposure. This may be either due to a different deposition of cigarette smoke in the lung due to differences in lung growth and development in males and females, a different inhalation pattern, or different response mechanisms to the inhaled toxic substances. Subsequently, other mechanisms than oxidative stress may play a more prominent role in COPD development in women, such as airway remodeling by inflammatory cells, in interaction with proteins, proteolytic enzymes and cytokines.
The *GSTP1* Ile105Val SNP has previously been associated with COPD [14]. He *et al*. showed that *GSTP1* Ile105Val was associated with lower lung function and with excess FEV_1 decline in subjects who had a high baseline lung function at inclusion in the study [30]. We found no association of the *GSTP1* SNPs with lung function and COPD cross-sectionally or with longitudinal FEV_1 decline, also not in interaction with smoking or gender (data not shown). We can thus conclude from our study that the effect of the *GSTP1* SNPs is unlikely to contribute to FEV_1 decline at a general population level. This conclusion is supported by Imboden *et al*. who also did not find association of the *GSTP1* Ile105Val SNP and FEV_1 decline in their general population study.

In conclusion, we confirm that *GSTM1* and *GSTT1* null alleles are important predictors for excess lung function decline in the general population. We show that especially males are at increased risk for excess FEV_1 decline, which is further enhanced when they smoke. Further studies have to assess whether these sex differences are due to different underlying types of COPD with different pathologic substrates and pathophysiology

ACKNOWLEDGEMENTS

This study was funded by the Dutch Asthma Foundation (NAF 3.2.02.51).

REFERENCES

1. Rytila P, Rehn T, Ilumets H, *et al.* Increased oxidative stress in asymptomatic current chronic smokers and GOLD stage 0 COPD. *Respir Res* 2006;7:69.

2. Rahman I. Oxidative stress in pathogenesis of chronic obstructive pulmonary disease: cellular and molecular mechanisms. *Cell Biochem Biophys* 2005;43(1):167-88.

3. Rahman I, Morrison D, Donaldson K, *et al.* Systemic oxidative stress in asthma, COPD, and smokers. *Am J Respir Crit Care Med* 1996;154(4 Pt 1):1055-60.

4. Kanazawa H, Yoshikawa J. Elevated oxidative stress and reciprocal reduction of vascular endothelial growth factor levels with severity of COPD. *Chest* 2005;128(5):3191-7.

5. Rahman I, Marwick J, Kirkham P. Redox modulation of chromatin remodeling: impact on histone acetylation and deacetylation, NF-kappaB and pro-inflammatory gene expression. *Biochem Pharmacol* 2004;68(6):1255-67.

6. Cantin AM, North SL, Hubbard RC, *et al.* Normal alveolar epithelial lining fluid contains high levels of glutathione. *J Appl Physiol* 1987;63(1):152-7.

7. Sherratt PJ, Pulford DJ, Harrison DJ, *et al.* Evidence that human class Theta glutathione S-transferase T1-1 can catalyse the activation of dichloromethane, a liver and lung carcinogen in the mouse. Comparison of the tissue distribution of GST T1-1 with that of classes Alpha, Mu and Pi GST in human. *Biochem J* 1997;326 (Pt 3):837-46.

8. Seidegard J, Pero RW, Miller DG, *et al.* A glutathione transferase in human leukocytes as a marker for the susceptibility to lung cancer. *Carcinogenesis* 1986;7(5):751-3.

9. Baranova H, Perriot J, Albuisson E, *et al.* Peculiarities of the GSTM1 0/0 genotype in French heavy smokers with various types of chronic bronchitis. *Hum Genet* 1997;99(6):822-6.

10. Cheng SL, Yu CJ, Chen CJ, *et al.* Genetic polymorphism of epoxide hydrolase and glutathione S-transferase in COPD. *Eur Respir J* 2004;23(6):818-24.

11. Gilliland FD, Gauderman WJ, Vora H, *et al.* Effects of glutathione-S-transferase M1, T1, and P1 on childhood lung function growth. *Am J Respir Crit Care Med* 2002;166(5):710-6.

12. Imboden M, Downs SH, Senn O, *et al.* Glutathione S-transferase genotypes modify lung function decline in the general population: SAPALDIA cohort study. *Respir Res* 2007;8:2.

13. Sundberg K, Johansson AS, Stenberg G, *et al.* Differences in the catalytic efficiencies of allelic variants of glutathione transferase P1-1 towards carcinogenic diol epoxides of polycyclic aromatic hydrocarbons. *Carcinogenesis* 1998;19(3):433-6.

14. Ishii T, Matsuse T, Teramoto S, *et al.* Glutathione S-transferase P1 (GSTP1) polymorphism in patients with chronic obstructive pulmonary disease. *Thorax* 1999;54(8):693-6.

15. Vibhuti A, Arif E, Deepak D, S *et al.* Genetic polymorphisms of GSTP1 and mEPHX correlate with oxidative stress markers and lung function in COPD. *Biochem Biophys Res Commun* 2007;359(1):136-42.

16. van Diemen CC, Postma DS, Vonk JM, *et al.* A disintegrin and metalloprotease 33 polymorphisms and lung function decline in the general population. *Am J Respir Crit Care Med* 2005;172(3):329-33.

17. van Diemen CC, Postma DS, Vonk JM, B *et al.*Decorin and TGF-beta1 polymorphisms and development of COPD in a general population. *Respir Res* 2006;7:89.

18. Fabbri LM, Hurd SS. Global Strategy for the Diagnosis, Management and Prevention of COPD: 2003 update. *Eur Respir J* 2003;22(1):1-2.

19. Pinheiro JC, Bates DM. Mixed-Effects Models in S and S-Plus. New York, NY: Springer; 2000.

20. R development Core Team. R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing; 2004.

21. Actis-Goretta L, Carrasquedo F, Fraga CG. The regular supplementation with an antioxidant mixture decreases oxidative stress in healthy humans. Gender effect. *Clin Chim Acta* 2004;349(1-2):97-103.

22. Tujague J, Bastaki M, Holland N, *et al*. Antioxidant intake, GSTM1 polymorphism and pulmonary function in healthy young adults. *Eur Respir J* 2006;27(2):282-8.

23. Grievink L, Smit HA, Ocke MC, *et al*. Dietary intake of antioxidant (pro)-vitamins, respiratory symptoms and pulmonary function: the MORGEN study. *Thorax* 1998;53(3):166-71.

24. Ochs-Balcom HM, Grant BJ, Muti P, *et al*. Antioxidants, oxidative stress, and pulmonary function in individuals diagnosed with asthma or COPD. *Eur J Clin Nutr* 2006;60(8):991-9.

25. Mucha L, Stephenson J, Morandi N, *et al*. Meta-analysis of disease risk associated with smoking, by gender and intensity of smoking. *Gen Med* 2006;3(4):279-91.

26. Gan WQ, Man SF, Postma DS, *et al*. Female smokers beyond the perimenopausal period are at increased risk of chronic obstructive pulmonary disease: a systematic review and meta-analysis. *Respir Res* 2006;7:52.

27. Watson L, Vonk JM, Lofdahl CG, *et al*. Predictors of lung function and its decline in mild to moderate COPD in association with gender: results from the Euroscop study. *Respir Med* 2006;100(4):746-53.

28. Dransfield MT, Washko GR, Foreman MG, *et al*. Gender differences in the severity of CT emphysema in COPD. *Chest* 2007;132(2):464-70.

29. Martinez FJ, Curtis JL, Sciurba F, *et al*. Gender Differences in Severe Pulmonary Emphysema. *Am J Respir Crit Care Med* 2007;176(3):243-52.

30. He JQ, Ruan J, Connett JE, *et al*. Antioxidant gene polymorphisms and susceptibility to a rapid decline in lung function in smokers. *Am J Respir Crit Care Med* 2002;166(3):323-8.

CHAPTER 9

Superoxide Dismutases variants, lung function and bronchial responsiveness in the general population

Mateusz Siedliński
Cleo C van Diemen
Dirkje S Postma
Judith M Vonk
H Marike Boezen

Submitted for publication

ABSTRACT

Background

Oxidative stress is an important causative factor in the onset and progression of smoking-related lung diseases like chronic obstructive pulmonary disease (COPD). Superoxide dismutases (SODs) prevent the elevation of oxidative stress by degradation of superoxide anions. Therefore we aimed to investigate the association of single nucleotide polymorphisms (SNPs) located in *SODs* (*SOD-2* and *SOD-3*) with development of bronchial hyperresponsiveness (BHR) and COPD, and the longitudinal course of forced expiratory volume in 1 second (FEV_1) over time.

Methods

1390 subjects from the prospective Vlagtwedde-Vlaardingen cohort were genotyped for 2 haplotype-tagging SNPs in *SOD-2* and 3 nonsynonymous SNPs in *SOD-3*. COPD was defined as GOLD stage ≥ II at the last survey. BHR was expressed dichotomously as a $PC_{10} \leq 8$ mg/ml of histamine. Presence of COPD and/or BHR was analyzed using binary and multinomial logistic regression analyses. FEV_1 decline was analyzed using linear mixed effect models.

Results

The C7693T SNP, located in intron 3 of *SOD-2,* was significantly associated with the presence of COPD and BHR in the total population. Multinomial logistic regression analyses showed that the T/T genotype for this SNP and the Val/Val genotype for the *SOD-2* Ala16Val substitution were risk factors for BHR in individuals without COPD. *SOD-3* Arg213Gly substitution was associated with slower FEV_1 decline in never-smokers exclusively.

Conclusions

Both *SOD-2* SNPs are risk factors for BHR in absence of COPD while the *SOD-2* C7693T SNP is subsequently a risk factor for COPD in the total population. *SOD-3* Arg213Gly SNP is protective for FEV_1 decline in never-smokers.

INTRODUCTION

An imbalance between oxidants and antioxidants is considered to be an important causative factor in onset and progression of chronic obstructive pulmonary disease (COPD)[1, 2]. A key molecule in the regulatory process of oxidative stress is superoxide anion, which is inhaled with cigarette smoke and additionally generated during numerous cellular reactions [3]. Free superoxide anion is a substrate for the synthesis of highly reactive oxygen species which can damage epithelial cells, impair epithelial ciliary function, alter expression of cellular adhesion molecules, and increase airway smooth muscle contraction in response to histamine and other stimuli in *in vitro* and in *in vivo* models [4]. Therefore, impaired superoxide metabolism may contribute to the development of bronchial hyperresponsiveness (BHR) and the subsequent onset of COPD [5, 6]. Indeed we have previously shown that patients with COPD have increased production of superoxide anion by blood leukocytes, which was associated with more severe hyperresponsiveness [7]. However, this cross-sectional study did not elucidate whether the increased production of superoxide anion was the cause or consequence of BHR and COPD. So far, studies on the relation between superoxide anion and both the development of COPD and BHR in the general population are lacking.

The family of superoxide dismutases (SODs) is the sole unique enzymatic system able to degrade a superoxide anion [8]. Impairment of the mitochondrial superoxide dismutase (SOD-2) activity was related to asthma pathophysiology [9] and extracellular superoxide dismutase (SOD-3) was shown to protect lungs against oxidant-mediated injury [10, 11]. Therefore it is of interest to study whether Single Nucleotide Polymorphisms (SNPs) in these *SOD* genes contribute to the development of BHR and/or COPD. *SOD-2* contains one prevalent nonsynonymous SNP (nsSNP) i.e. Ala16Val which is localized in a protein signal sequence, thus possessing a possibly functional role [12, 13]. This substitution has been associated with a higher lung cancer risk in Caucasians [14-16], but not with the presence of COPD in Caucasian and Chinese smokers [17, 18]. *SOD-3* contains three nsSNPs i.e. Ala40Thr, Phe131Cys and Arg213Gly. The Arg213Gly substitution has been associated with elevated SOD-3 levels in human plasma [19] and with protection against COPD development in Caucasians [17, 20], whereas Ala40Thr and Phe131Cys have not yet been studied in this context.

This study links the well-known risk factors for COPD development, i.e. BHR and smoking, with polymorphisms located in genes involved in response to oxidative stress, a key manifestation in COPD pathogenesis. The study was performed in a population-based cohort that was prospectively followed for 25 years with lung function measurements every 3 years.

METHODS

Subjects

A total of 1390 subjects of the Vlagtwedde-Vlaardingen cohort who participated in the last survey (1989–1990) were included in the study. This general population-based cohort of Caucasian individuals, of Dutch descent, started in 1965 and has been followed-up for 25 years. Surveys were performed at 3 year intervals. Details on the cohort have been described previously (*see* also the online supplement for details) [21-23]. The study protocol was approved by the local university hospital medical ethics committee and all participants gave their written informed consent.

DNA extraction and genotyping

See the online supplement for a description of DNA extraction and the genotyping protocol. Five SNPs in two *SODs* genes were genotyped, i.e.: Ala16Val (rs4880) and C7693T (rs2842958) in *SOD-2* and Ala40Thr (rs2536512), Arg213Gly (rs1799895) and Phe131Cys [24] in *SOD-3*. The C7693T SNP (see AY267901 sequence in GenBank®) was included to be able to tag *SOD-2* haplotypes, as described previously [25].

Presence of COPD and bronchial hyperresponsiveness (BHR) phenotype

We identified subjects with COPD according to the Global Initiative for Chronic Obstructive Lung Disease (GOLD) criteria, i.e. an $FEV_1/VC < 70\%$ and an $FEV_1 < 80\%$ predicted (GOLD stage II or higher) [26] at the last survey. Bronchial responsiveness to histamine was assessed in the random subsample of the total cohort as described previously (30 seconds method) [23]. BHR at the last survey was expressed dichotomously (as a $PC_{10} \leq 8$ mg/ml histamine).

Statistics

Differences in prevalence of rare alleles of SNPs between subjects with and without COPD or BHR were tested with X^2 test and logistic regression. Additionally, we performed all mentioned analyses stratified according to smoking habits (never- and ever-smokers), since we expected the SODs effects to depend on smoking habits.

In order to disentangle the phenotypes COPD and BHR, subjects were classified into 4 phenotypic groups according to the presence of COPD and/or BHR (i.e.: 'COPD-/BHR-', 'COPD-/BHR+', 'COPD+/BHR-' and 'COPD+/BHR+'). Multinomial logistic regression, adjusted for packyears of smoking, was performed for all SNPs separately, with the 4 groups as dependent variable. These analyses allowed us to predict if subjects with certain genotypes were more likely to develop a COPD phenotype or a BHR phenotype.

Statistical analyses were performed using SPSS (version 14.0.1 for Windows; SPSS, Chicago, IL). LME models from the statistical package R (version 1.9.1) [27] were used to estimate longitudinal changes in lung function in the total cohort as well as in the ever-

and never-smokers subgroups. PHASE (version 2.1) [28, 29] was used to estimate *SODs* haplotype frequency. P values < 0.05 were considered to be significant.

RESULTS

Study population and distribution of *SODs* SNPs

The study population characteristics and genotype frequencies are shown in table 1. All SNPs were distributed according to Hardy Weinberg Equilibrium in the total population. The Phe131Cys substitution in *SOD-3* was not present in our cohort. We found no homozygotes mutants for the Arg213Gly SNP in *SOD-3* in our population.

Table 1: Characteristics of the sample of the Vlagtwedde–Vlaardingen cohort at the last survey

	Total population n=1390	Ever-smokers* n=945
Males, n (%)	714 (51.4)	610 (64.6)
Age in years, median (range)	52 (35–79)	50 (35-79)
Pack-years of smoking, median (range)	8.0 (0–262.2)	18.8 (0-262.2)
FEV_1%predicted, mean (SD)	92.6 (15.3)	90.6 (15.5)
FEV_1/VC %, mean (SD)	73.9 (8.7)	72.4 (9.2)
COPD – GOLD stage II or higher, n (%)	167 (12.4)	140 (15.3)
BHR – PC_{10}≤8 mg/ml histamine, n (%)[§]	190 (46.5)	141 (50.4)
Genotype frequency (%heterozygotes / %homozygotes mutant):		
SOD-2 Ala16Val:	49.6/23.8	49.8/23.6
SOD-2 C7693T:	29.3/4.5	28.8/5.2
SOD-3 Ala40Thr:	43.5/11.9	42.0/12.1
SOD-3 Arg213Gly:	3.8/0.0	4.1/0.0

*sub-population of the total population

[§]BHR test performed on a random subgroup (n=409) of the total population (n=1390)

Abbreviations: SD Standard Deviation; BHR Bronchial Hyperresponsiveness; COPD Chronic Obstructive Pulmonary Disease; GOLD Global Initiative for Chronic Obstructive Lung Disease; FEV_1 Forced Expiratory Volume in one second; VC Vital Capacity; PC_{10} = the histamine concentration causing an FEV_1 decrease of 10% or more from baseline

SODs SNPs and the presence of COPD

The *SOD-2* C7693T SNP was significantly associated with COPD in the total population and in ever-smokers (table 2 and table S2 in supplementary data), but not in never-smokers (data not shown). No significant associations were found for the other investigated SNPs (table 2).

SODs SNPs and the presence of BHR

The *SOD-2* C7693T SNP was significantly associated with the presence of BHR within the total population (table 2 and table S3 in supplementary data). Mutant homozygotes had a significantly higher risk of having BHR compared to wild types. The Ala16Val substitution

in *SOD-2* was borderline significantly associated with BHR (p=0.07 for mutant vs. wild type homozygotes, table 2). Stratified analyses according to smoking showed no significant associations for these two SNPs. The Ala40Thr SNP was significantly associated with BHR only in never-smokers (OR [95% C.I.] for the presence of BHR: Ala/Thr vs. Ala/Ala 2.84 [1.08-7.48]; and Thr/Thr vs. Ala/Ala 14.40 [2.92-71.08] adjusted for packyears of smoking and FEV_1% predicted, see also table S4 in supplementary data), but not in the total population and in ever-smokers (table 2).

Table 2: Odds ratios (95% C.I.) for the presence of COPD respectively BHR, in the total population and in ever-smokers, according to *SODs* SNPs Logistic regression analysis with OR for COPD adjusted for pack-years, and OR for BHR adjusted for pack-years and FEV_1% predicted

Gene	SNP	MAF	OR [C.I.] (heterozygotes vs. wild type)	OR [C.I.] (homozygotes vs. wild type)
			COPD in total population (n=1390)	
SOD-2	Ala16Val	0.49	1.05 [0.66-1.67]	0.94 [0.62-1.41]
	C7693T	0.21	0.74 [0.50-1.11]	**1.98 [1.04-3.78]***
SOD-3	Ala40Thr	0.34	1.03 [0.72-1.49]	0.59 [0.31-1.13]
	Arg213Gly	0.02	1.66 [0.77-3.58]	-
			COPD in ever-smokers (n=945)	
SOD-2	Ala16Val	0.48	0.93 [0.59-1.46]	1.19 [0.71-1.98]
	C7693T	0.21	0.75 [0.48-1.12]	**2.37 [1.20-4.67]***
SOD-3	Ala40Thr	0.33	1.18 [0.70-1.78]	0.68 [0.34-1.36]
	Arg213Gly	0.02	1.96 [0.89-4.36]	-
			BHR in total population (n=409)	
SOD-2	Ala16Val	0.45	1.24 [0.74-2.08]	1.80 [0.97-3.35]¥
	C7693T	0.20	1.14 [0.61-2.13]	**3.21 [1.03-10.03]***
SOD-3	Ala40Thr	0.32	1.01 [0.63-1.62]	1.31 [0.63-2.71]
	Arg213Gly	0.01	0.33 [0.07-1.52]	-
			BHR in ever-smokers (n=280)	
SOD-2	Ala16Val	0.46	1.62 [0.87-3.03]	1.61 [0.76-3.41]
	C7693T	0.19	0.95 [0.54-1.69]	2.97 [0.81-10.98]§
SOD-3	Ala40Thr	0.32	0.67 [0.38-1.21]	0.57 [0.24-1.36]
	Arg213Gly	0.01	0.48 [0.09-2.64]	-

*p<0.05 ¥p=0.07 §p=0.10

Abbreviations: BHR Bronchial Hyperresponsiveness; SNP Single Nucleotide Polymorphism; MAF Minor Allele Frequency in the investigated group; OR Odds Ratio ; C.I. Confidence Intervals; COPD Chronic Obstructive Pulmonary Disease; SOD Superoxide Dismutase

SODs SNPs and the presence of COPD and/or BHR phenotype

Two hundred and thirteen subjects had neither COPD nor BHR (COPD-/BHR-), 154 subjects had COPD-/BHR+, 6 subjects COPD+/BHR-, and 36 subjects had COPD+/BHR+. Subjects with the homozygous Ala16Val substitution in *SOD-2* were most likely to be COPD-/BHR+ (OR=2.1; 95% C.I.: 1.1-3.8 for the Val/Val variant compared to Ala/Ala), indicating that the Ala16Val substitution was predictive for the presence of BHR and not for COPD or their combination. Likewise, subjects with the homozygous mutant *SOD-2* C7693T SNP were most likely to have COPD-/BHR+ (OR=3.1; 95% C.I.: 1.03-9.51 for the homozygous mutant

variant compared to the wild type). The polymorphisms in the *SOD-3* gene were not associated with any of the 4 phenotypic outcome groups.

SODs SNPs and longitudinal change in FEV_1

The annual FEV_1 decline was similar in all genotype subgroups of the investigated SNPs (see table S5 in the online supplement). Stratification according to smoking habits revealed that the *SOD-3* Arg213Gly substitution was associated with slower FEV_1 decline in never-smokers, e.g. 9.3 ml less decline per year in carriers of the Arg/Gly genotype while compared to the wild type (figure 1). The other SNPs were not associated with excess FEV_1 decline in stratified analyses.

SOD-2 and *SOD-3* haplotypes

Within *SOD-3* three major haplotypes existed i.e.: Ala-Arg (frequency 66%), Thr-Arg (32%) and Thr-Gly (2%) (the first amino acid corresponds to the Ala40Thr substitution, the second to the Arg213Gly substitution).

We identified three major *SOD-2* haplotypes in the total population, i.e. CC (frequency 51%), TC (28%) and TT (21%), which is in agreement with another study performed in a Caucasian population [25]. The mutant allele of C7693T SNP (i.e. "T" allele) coexisted only with the mutant variant of the Ala16Val substitution (i.e. "T" allele corresponding to the valine), and consequently, is present only in the **TT** *SOD-2* haplotype. Therefore, the associations found with the "T" allele of the C7693T SNP can represent the effect of the **TT** haplotype as well.

Discussion

Of all human tissues, the lungs are most directly and particularly exposed to noxious free radicals. Tobacco smoke and environmental air pollution are the major sources of these particles. Therefore, the redox balance regulation can be an important factor for the development of bronchial hyperresponsiveness, excess decline in lung function, and development of COPD. We showed that *SOD-2*, containing the mutant allele of the C7693T SNP, is a risk factor for bronchial hyperresponsiveness. In the general population and exclusively in subjects who have smoked, COPD was more prevalent in carriers of two copies of the mutant allele compared to wild type carriers. The *SOD-3* Arg213Gly substitution, that has previously been shown to be protective for COPD [17, 20], was associated with slower FEV_1 decline in never-smokers only.

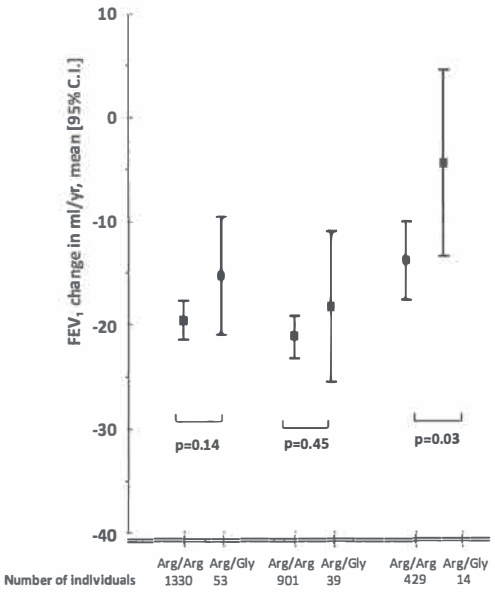


Figure 1: Adjusted means for the annual change in FEV_1 for the SOD-3 Arg213Gly genotypes in the total population and according to smoking habits
FEV_1 Forced Expiratory Volume in one second; SOD Superoxide Dismutase

Our finding is of interest since SOD-2 dysfunction and inactivation has been recently described *ex vivo* in bronchial epithelial cells derived from airways of asthmatic patients [9]. This confirms the importance of SOD-2 in lung homeostasis where reactive oxygen species can impair epithelial ciliary function and increase airway smooth muscle contraction in response to histamine [4], which may contribute to the development of BHR. Our data suggest that this may also be the case in COPD.

We found the mutant homozygosity of the *SOD-2* SNP (C7693T), located in intron 3, to be a risk factor for the presence of COPD and BHR in the total population at the last survey. Although the role of SNPs localized in introns is not clear yet, it is known that they may affect alternative splicing process [30, 31] which result in different amino acid sequence in mature protein. The main difference between the *SOD-2* isoforms is the presence of an additional exon within the third intron [32], but the implication with respect to putative differences in activity between both isoforms is yet unknown. However, it is indeed confirmed that SOD-2 isoforms formation may depend on genetic variations which has been *in vitro* shown for a polymorphism (single base pair deletion) that is prevalent in African-Americans [33]. This study concluded however that other (cis-acting) factors are necessary for switching between two isoforms *in vivo*. Thus it is also possible that C7693T SNP is in high linkage disequilibrium with another (functional) SNP, which is actually responsible for the associations that we observed.

Theoretically, the Ala16Val substitution could be of importance as well, since *in vitro* studies have shown impaired transport of Val16-SOD-2 to the mitochondrion [12, 13], a major cellular source of reactive oxygen species. However, we did not find an association of the Ala16Val SNP with either the presence of COPD or with lung function decline. This finding is in concordance with results of two previous studies performed by Young *et al.* [17] and Mak *et al.* [18], who found no association of the Ala16Val with COPD in case-control studies (smokers with COPD versus non-susceptible smokers).

Bronchial responsiveness may be modulated by superoxide anion production [7] and BHR is a known risk factor for COPD development [5, 6]. Due to these facts, impaired superoxide radical detoxification, caused by loss of SOD-2 function, may increase the level of BHR and affect the development of COPD. Since it is difficult to disentangle the BHR and the COPD phenotype, we performed multinomial logistic regression analyses to estimate whether certain genotypes are more likely to lead to 'pure' BHR (with absence of COPD) or 'pure' COPD (with absence of BHR), or were likely to lead to both BHR and COPD, or neither. Remarkably, subjects with homozygous variants of the *SOD-2* Ala16Val substitution or the *SOD-2* C7693T SNP showed increased odds ratios (OR=2.1 and 3.7 respectively) for having 'pure BHR', which suggests it is indeed the phenotype BHR that is associated with these *SOD-2* SNPs, and the relation that we have found between C7693T SNP and COPD is driven by the association with BHR. We must, however, take into account that the power to identify genetic associations with the 'pure COPD' phenotype was limited due to the low number of BHR-/COPD+ subjects.

The *SOD-2* **TT** haplotype contains rare alleles of both investigated *SOD-2* SNPs. Moreover, it is the only haplotype that contains the mutant allele of C7693T SNP. Remarkably, we found an association of the homozygous mutant of C7693T with BHR, but not with the heterozygous mutant of C7693T. Therefore, we conclude that the effect of this SNP, and consequently the **TT** haplotype, is important in the recessive mode. Similarly, within ever-smokers and within the total population, C7693T homozygotes are at risk to develop COPD, whereas C7693T heterozygotes are not. Thus, we can also assume a recessive effect of this haplotype, and consequently of the C7693T SNP, with respect to COPD development.

Juul and colleagues [20] have shown prospectively and cross-sectionally a lower risk for COPD for the *SOD-3* Arg213Gly SNP heterozygotes while compared to the wild type. This has been confirmed by another cross-sectional study [17]. The underlying mechanism of these associations may be explained by results from *in vitro* studies [19, 34], which showed mutated SOD-3 protein to be released from heparan sulphate in the cellular membrane into the tissue interstitium. We additionally found that the protective, effect of this SNP on longitudinal FEV_1 change appears in never-smokers. Therefore we hypothesize that there are other antioxidant-related genetic factors that additionally contribute to the FEV_1 change in smokers. We must however admit that the low prevalence of this SNP might

have underpowered our study thus we were not able to detect possible associations neither with COPD in the general population nor with FEV_1 decline in ever-smokers. We found no associations of the Ala40Thr with either the presence of COPD or BHR in the total population. More detailed analysis showed that this SNP may play a role in BHR presence in never smokers. However, since this substitution in *SOD-3* is not yet examined for its functional implication, more research on its function is required.

In summary, we conclude that SOD-2 is an antioxidant enzyme in which variations in DNA sequence can play a role in the development of BHR. By showing the protective effect of the Arg213Gly SNP in the *SOD-3* on FEV_1 course in never-smokers we put another piece of puzzle to the knowledge on the role of this substitution in pulmonary disease. Further studies are needed to confirm the possible role of *SODs* SNPs in COPD and BHR, and further functional studies are warranted.

Funding

Dutch Asthma Foundation (grant 3.2.02.51), The Netherlands; University of Groningen, The Netherlands

REFERENCES

1. Kinnula VL. Focus on antioxidant enzymes and antioxidant strategies in smoking related airway diseases. *Thorax* 2005;60:693-700.
2. Domej W, Foldes-Papp Z, Flogel E, *et al*. Chronic obstructive pulmonary disease and oxidative stress. *Curr Pharm Biotechnol* 2006;7:117-123.
3. Tsuchiya M, Thompson DF, Suzuki YJ, *et al*. Superoxide formed from cigarette smoke impairs polymorphonuclear leukocyte active oxygen generation activity. *Arch Biochem Biophys* 1992;299:30-37.
4. Henricks PA, Nijkamp FP. Reactive oxygen species as mediators in asthma. *Pulm Pharmacol Ther* 2001;14:409-420.
5. Postma DS, Boezen HM. Rationale for the Dutch hypothesis. Allergy and airway hyperresponsiveness as genetic factors and their interaction with environment in the development of asthma and COPD. *Chest* 2004;126:96S-104S.
6. Brutsche MH, Downs SH, Schindler C, *et al*. Bronchial hyperresponsiveness and the development of asthma and COPD in asymptomatic individuals: SAPALDIA cohort study. *Thorax* 2006;61:671-677.
7. Postma DS, Renkema TE, Noordhoek JA, *et al*. Association between nonspecific bronchial hyperreactivity and superoxide anion production by polymorphonuclear leukocytes in chronic air-flow obstruction. *Am Rev Respir Dis* 1988;137:57-61.
8. Fridovich I. Superoxide anion radical (O2-.), superoxide dismutases, and related matters. *J Biol Chem* 1997;272:18515-18517.
9. Comhair SA, Xu W, Ghosh S, *et al*. Superoxide dismutase inactivation in pathophysiology of asthmatic airway remodeling and reactivity. *Am J Pathol* 2005;166:663-674.
10. Carlsson LM, Jonsson J, Edlund T, *et al*. Mice lacking extracellular superoxide dismutase are more sensitive to hyperoxia. *Proc Natl Acad Sci USA* 1995;92:6264-6268.
11. Folz RJ, Abushamaa AM, Suliman HB. Extracellular superoxide dismutase in the airways of transgenic mice reduces inflammation and attenuates lung toxicity following hyperoxia. *J Clin Invest* 1999;103:1055-1066.
12. Sutton A, Khoury H, Prip-Buus C, *et al*. The Ala16Val genetic dimorphism modulates the import of human manganese superoxide dismutase into rat liver mitochondria. *Pharmacogenetics* 2003;13:145-157.
13. Sutton A, Imbert A, Igoudjil A, *et al*. The manganese superoxide dismutase Ala16Val dimorphism modulates both mitochondrial import and mRNA stability. *Pharmacogenet Genomics* 2005;15:311-319.
14. Wang LI, Miller DP, Sai Y, *et al*. Manganese superoxide dismutase alanine-to-valine polymorphism at codon 16 and lung cancer risk. *J Natl Cancer Inst* 2001;93:1818-1821.
15. Wang LI, Neuberg D, Christiani DC. Asbestos exposure, manganese superoxide dismutase (MnSOD) genotype, and lung cancer risk. *J Occup Environ Med* 2004;46:556-564.
16. Liu G, Zhou W, Wang LI, *et al*. MPO and SOD2 polymorphisms, gender, and the risk of non-small cell lung carcinoma. *Cancer Lett* 2004;214:69-79.
17. Young RP, Hopkins R, Black PN, *et al*. Functional variants of antioxidant genes in smokers with COPD and in those with normal lung function. Thorax 2006;61:394-399.
18. Mak JC, Ho SP, Yu WC, *et al*. Polymorphisms and functional activity in SOD and catalase genes in smokers with COPD. *Eur Respir J* 2007.
19. Sandstrom J, Nilsson P, Karlsson K, *et al*. 10-fold increase in human plasma extracellular superoxide dismutase content caused by a mutation in heparin-binding domain. *J Biol Chem* 1994;269:19163-19166.

20. Juul K, Tybjaerg-Hansen A, Marklund S, *et al.* Genetically increased antioxidative protection and decreased chronic obstructive pulmonary disease. *Am J Respir Crit Care Med* 2006;173:858-864.

21. van Diemen CC, Postma DS, Vonk JM, *et al.* Decorin and TGF-beta1 polymorphisms and development of COPD in a general population. *Respir Res* 2006;7:89.

22. van Diemen CC, Postma DS, Vonk JM, *et al.* A disintegrin and metalloprotease 33 polymorphisms and lung function decline in the general population. *Am J Respir Crit Care Med* 2005;172:329-333.

23. Rijcken B, Schouten JP, Mensinga TT, *et al.* Factors associated with bronchial responsiveness to histamine in a population sample of adults. Am Rev Respir Dis 1993;147:1447-1453.

24. Campo S, Sardo AM, Campo GM, *et al.* Extracellular superoxide dismutase (EC-SOD) gene mutations screening in a sample of Mediterranean population. *Mutat Res* 2005;578:143-148.

25. Cebrian A, Pharoah PD, Ahmed S, *et al.* Tagging single-nucleotide polymorphisms in antioxidant defense enzymes and susceptibility to breast cancer. *Cancer Res* 2006;66:1225-1233.

26. From the Global Strategy for the Diagnosis, Management and Prevention of COPD, Global Initiative for Chronic Obstructive Lung Disease (GOLD) 2006. Available from: http://www.goldcopd.org.

27. R Development Core Team. R: a language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing. 2004.

28. Stephens M, Smith NJ, Donnelly P. A new statistical method for haplotype reconstruction from population data. *Am J Hum Genet* 2001;68:978-989.

29. Stephens M, Scheet P. Accounting for decay of linkage disequilibrium in haplotype inference and missing-data imputation. *Am J Hum Genet* 2005;76:449-462.

30. von AN, Oellerich M. The intronic prothrombin 19911A>G polymorphism influences splicing efficiency and modulates effects of the 20210G>A polymorphism on mRNA amount and expression in a stable reporter gene assay system. *Blood* 2004;103:586-593.

31. Kuehl P, Zhang J, Lin Y, *et al.* Sequence diversity in CYP3A promoters and characterization of the genetic basis of polymorphic CYP3A5 expression. *Nat Genet* 2001;27:383-391.

32. Hubbard T, Barker D, Birney E, *et al.* The Ensembl genome database project. *Nucleic Acids Res* 2002;30:38-41.

33. Shao J, Chen L, Marrs B, *et al.* SOD2 polymorphisms: unmasking the effect of polymorphism on splicing. *BMC Med Genet* 2007;8:7.

34. Folz RJ, Peno-Green L, Crapo JD. Identification of a homozygous missense mutation (Arg to Gly) in the critical binding region of the human EC-SOD gene (SOD3) and its association with dramatically increased serum enzyme levels. *Hum Mol Genet* 1994;3:2251-2254.

Online Data Supplement to paper:

Superoxide Dismutases variants, lung function and bronchial hyperresponsiveness in the general population

Mateusz Siedliński, Cleo C van Diemen, Dirkje S Postma,
Judith M Vonk, H Marike Boezen

METHODS

Subjects

From all subjects information was collected on respiratory symptoms, smoking status, age, and sex by the Dutch version of the British Medical Council standardized questionnaire. Spirometry was performed using a standardized protocol [1].
Subjects with forced expiratory volume (FEV_1) lower than 1.5 L and those who could not perform a forced expiratory manoeuvre were excluded from responsiveness testing, as were subjects suffering from heart disease, hypertension, or acute respiratory infections.

DNA extraction and genotyping

Neutrophil depots from centrifuged blood samples were collected and stored at −20°C. DNA was extracted with a QIAamp DNA blood mini kit (Qiagen, Hilden, Germany) and checked for purity and concentration with a NanoDrop ND-1000 UV–Vis spectrophotometer (NanoDrop Technologies, Wilmington, DE). We genotyped DNA samples of those subjects with more than 1,500 ng of isolated DNA available (n = 1,390). Genotyping was performed using Applied Biosystems TaqMan® SNP Genotyping Assays (Nieuwekerk aan de IJssel, The Netherlands) as described previously [1]. Sequences of primers and probes used for genotyping are shown in table S1.Due to the low fluorescence signal at the end-point measurement and in order to ensure genotyping quality, SNP rs2842958 in *SOD-2* was genotyped two times. Subjects with conflicting (n=18) genotypes for both independent genotyping rounds were excluded from the further analyses.

Statistics

Linear Mixed Effects (LME) models were used to investigate the effect of SOD SNPs on the annual decline in FEV_1. Time was defined as the time in years relative to the first FEV_1 measurement and corresponding age. FEV_1 measurements were included from the age of 30 years, because an individual's maximal achieved lung function is assumed to have been reached before that age and lung function is considered to be either in the plateau or decline phase [2]. Variables included in the model were sex, packyears of smoking, and the first available FEV_1 after age 30 years and their interaction with time.

TABLES

Table S1: Sequences of primers and probes

SNP	Primers	Probe 1*	Probe 2*
SOD-2 Ala16Val rs4880	Applied Biosystems assay ID: C___8709053_10		
SOD-2 C7693T rs2842958	**Forward:** ATGCCTGTAATCCCAGCTACTTG **Reverse:** CCTCCGCCTTTCAGGTTCAT	CTGAGACAC<u>A</u>AGAATT	TGAGACAC<u>G</u>AGAATT
SOD-3 Ala58Thr rs2536512	**Forward:** CTGGCAGGAGGTCATGCA **Reverse:** CCGACGGCTGCACCT	ACGACGACGGC<u>A</u>CG	CGACGACGGC<u>G</u>CG
SOD-3 Phe149Cys	**Forward:** GCACCCGGGCGACT **Reverse:** CCGGCGCGGTACCT	ACCGCG<u>A</u>AGTTGCCGA	CGCG<u>C</u>AGTTGCCGA
SOD-3 Arg213Gly rs1799895	Applied Biosystems assay ID: C___2307506_10		

*Underlined bases indicate specific polymorphic site recognized by each probe

Table S2: COPD prevalence in selected subgroups

SNP	MAF	Genotype	COPD total population, n [%][§]	P (2 df)*	COPD ever-smokers, n [%][§]	P (2 df)*
SOD-2						
rs4880, Ala16Val	0.49	Ala/Ala	44 [12.5]	0.74	36 [15.1]	0.41
		Ala/Val	75 [11.7]		61 [13.9]	
		Val/Val	42 [13.4]		38 [17.9]	
rs2842958, C7693T	0.21	C/C	110 [12.9]	0.01	92 [15.6]	0.008
		C/T	38 [9.6]		31 [12.1]	
		T/T	14 [23.0]		14 [29.8]	
SOD-3						
rs2536512, Ala58Thr	0.34	Ala/Ala	72 [12.8]	0.27	58 [14.6]	0.34
		Ala/Thr	68 [12.4]		58 [16.2]	
		Thr/Thr	12 [8.0]		11 [10.5]	
rs1799895,Arg213Gly	0.02	Arg/Arg	154 [12.0]	0.20	128 [14.7]	0.17
		Arg/Gly	9 [17.6]		9 [23.1]	

[§] within each genotype *in X^2 tests with all genotypes (2 df)

Abbreviations: MAF Minor Allele Frequency in the investigated group; SNP Single Nucleotide polymorphism; ; SOD Superoxide Dismutase; COPD Chronic Obstructive Pulmonary Disease

Table S3: BHR prevalence in selected subgroups

SNP	MAF	Genotype	BHR total population, n [%]§	P (2 df)*	BHR ever-smokers, n [%]§	P (2 df)*
SOD-2						
rs4880, Ala16Val	0.45	Ala/Ala	48 [40.0]	0.16	34 [42.5]	0.21
		Ala/Val	88 [46.8]		72 [54.5]	
		Val/Val	47 [53.4]		32 [53.3]	
rs2842958, C7693T	0.20	C/C	119 [44.6]	0.07	93 [49.7]	0.29
		C/T	52 [48.1]		34 [50.7]	
		T/T	13 [72.2]		10 [71.4]	
SOD-3						
rs2536512, Ala40Thr	0.32	Ala/Ala	84 [46.9]	0.89	70 [54.7]	0.33
		Ala/Thr	74 [45.1]		51 [49.0]	
		Thr/Thr	21 [48.8]		13 [40.6]	
s1799895, Arg213Gly	0.01	Arg/Arg	184 [46.7]	0.23	135 [50.2]	0.72
		Arg/Gly	3 [27.3]		3 [42.9]	

§ within each genotype *in X^2 tests with all genotypes (2 df)

Abbreviations: MAF Minor Allele Frequency in the investigated group; SOD Superoxide Dismutase; BHR Bronchial Hyperresponsiveness; SNP Single Nucleotide Polymorphism

Table S4: BHR prevalence at the last survey in never-smokers according to Ala40Thr *SOD-3* SNP

SOD-3 Ala40Thr genotype	BHR never-smokers, n [%]§	P (2 df)
Ala/Ala	14 [27.5]	0.02
Ala/Thr	23 [38.3]	
Thr/Thr	8 [72.7]	

§ within each genotype *in X^2 tests with all genotypes (2 df)

Abbreviations: SOD Superoxide Dismutase; BHR Bronchial Hyperresponsiveness; SNP Single Nucleotide Polymorphism

Table S5: LME models analysis on FEV_1 decline

SNP	MAF	Excess annual FEV_1 decline for heterozygotes, ml/yr [95% C.I.]*	P	Excess annual FEV_1 decline for homozygotes mutant ml/yr [95% C.I.]*	P
SOD-2					
rs4880 Ala16Val	0.49	-0.96 [-3.26 – 1.34]	0.42	-0.23 [-2.93 - 2.47]	0.87
rs2842958 C7693T	0.21	0.91 [-1.22 - 3.04]	0.40	2.66 [-1.84 - 7.16]	0.25
SOD-3					
rs2536512 Ala40Thr	0.34	0.05 [-2.05 - 2.15]	0.96	0.91 [-2.96 – 4.16]	0.58
rs1799895 Arg213Gly	0.02	4.28 [-1.40 - 9.96]	0.14	-	-

*compared to homozygotes wild type (reference); negative values indicate faster FEV_1 decline

Abbreviations: LME Linear Mixed Effect; MAF Minor Allele Frequency; SOD Superoxide Dismutase; FEV_1 Forced Expiratory Volume in one second; SNP Single Nucleotide Polymorphism; C.I. Confidence Interval

REFERENCES

1. van Diemen CC, Postma DS, Vonk JM, *et al.* A disintegrin and metalloprotease 33 polymorphisms and lung function decline in the general population. *Am J Respir Crit Care Med* 2005;172:329-333.

2. Rijcken B, Weiss ST. Longitudinal analyses of airway responsiveness and pulmonary function decline. *Am J Respir Crit Care Med* 1996;154:S246-S249.

CHAPTER 10

Pulmonary surfactant: genetic and lipidomic risk factors for Chronic Obstructive Pulmonary Disease

Cleo C van Diemen
Dirkje S Postma
Begona Barroso
Judith M Vonk
Rainer Bischoff
H Marike Boezen

Submitted for publication

ABSTRACT

Background

Pulmonary surfactant maintains alveolar structures, improves bronchial clearance, acts in innate immunity and attracts eosinophils. Therefore, changes in surfactant quantity or composition may contribute to lung function decline and development of Chronic Obstructive Pulmonary Disease (COPD).

Aim

To investigate 1) associations of single nucleotide polymorphisms (SNPs) in *surfactant protein* (*SFTP*) genes with decline in Forced Expiratory Volume in 1 second (FEV_1) and COPD development, and on well-known risk factors of COPD, i.e. bronchial hyperresponsiveness (BHR) and blood eosinophilia; 2) surfactant lipids and their degradation products in sputum of COPD patients and controls and their association with lung function.

Methods

SFTP genetics: we used longitudinal data from a cohort of the Dutch general population (follow-up 25 years, n=1,390). We estimated effects of SNPs in *SFTPA1*, *SFTPA2*, *SFTPB* and *SFTPD* on FEV_1 decline, COPD development, BHR and blood eosinophilia, adjusted for potential confounders. Lipid profiling: we analyzed surfactant phospholipids in sputum from 17 ex-smoking COPD patients and 11 ex-smoking healthy controls.

Results

SNPs *SFTPA1* Leu50Val and *SFTPD* Met11Thr were significant predictors of excess FEV_1 decline and COPD development. *SFTPD* Met11Thr was associated with blood eosinophilia. COPD patients had significantly lower phosphatidylcholine and higher lysophosphatidylcholine levels than controls. Accordingly, phosphatidylcholine levels were positively and lysophosphatidylcholine levels were negatively associated with lung function parameters.

Discussion

SFTP SNPs are associated with blood eosinophilia, a risk factor for accelerated FEV_1 decline. We show for the first time that genetic factors and lipid composition in surfactant can negatively affect lung function, thereby contributing to COPD development.

INTRODUCTION

Pulmonary surfactant covers all airway surfaces and alveolar spaces. Its major role is to decrease surface tension at the air–liquid interface, thereby reducing the tendency of alveoli to collapse during expiration. It also improves bronchial clearance and functions as part of the innate immune response against inhaled infectious agents and exposure to environmental pollutants.

The surfactant layer is produced by type II pneumocytes and consists of a complex mixture of lipids (about 90%) and specific surfactant proteins (SPs) (10%). The lipids consist for 80–90% of phospholipids and for 10–20% of cholesterol, triglycerides and free fatty acids. The principal phospholipids are phosphatidylcholine (PC) species (approximately 85%) containing high amounts of saturated palmitic acid (C16:0) and phosphatidylglycerol. Phospholipids can be degraded to lysophospholipids (LP) by loss of one of the fatty acids; lysophosphatidylcholines (LPC) are the most abundantly present LP species [1].
The main SPs are SP-A, B, C, and D (coded by the genes *SFTPA1*, *SFTPA2*, *SFTPB*, *SFTPC*, and *SFTPD)*. SP-A and D are hydrophilic, while SP-B and C are hydrophobic. The SPs facilitate the phospholipid arrangement in the lining of the alveoli to optimize their surface-tension reducing properties.

There is suggestive evidence that surfactant dysfunction plays a role in the development of Chronic Obstructive Pulmonary Disease (COPD), and this may be genetically determined. The *SFTP* genes are very good candidate genes for COPD development, although they have not been studied thoroughly in genetic research of COPD. The emphasis in this field of research has been on genes involved in the disturbed protease-antiprotease and oxidant-antioxidant balance in COPD. Single Nucleotide Polymorphisms (SNPs) in *SFTPA*, *SFTPB* and *SFTPD* have been associated previously with respiratory distress syndrome, COPD, and severe respiratory infections [2-6]. A lack in this field is a prospective cohort study to analyze the effects of SNPs in *SFTPs* on decline in lung function and development of COPD.

COPD is a disease characterized by chronic airway inflammation, airway obstruction and tissue remodeling, often accompanied by bronchial hyperresponsiveness (BHR).[7] We have previously shown that the presence of BHR in the general population predicts the onset of COPD-like symptoms such as chronic cough, phlegm and bronchitis, especially if eosinophilia is present [8]. Additionally, blood eosinophilia itself is an independent predictor of reduced lung function levels and accelerated lung function decline and indicates worse prognosis in chronic bronchitis [9, 10, 11]. Interestingly, SP-D regulates chemotaxis and degranulation of eosinophils which may connect the surfactant proteins, eosinophilia and reduced lung function [12, 13]. In the current study, we for the first time investigated the role of *SFTP* variants on lung function decline, BHR, blood eosinophilia, and development of COPD in a general population based cohort.

Surfactant dysfunction may also be caused by altered levels of surfactant phospholipids. Smoking is the major risk factor for COPD, and smokers have decreased phospholipid levels in bronchoalveolar lavage fluid (BALF) and an impaired surface activity of surfactants recovered from BALF [14,15]. In this study, we analyzed surfactant lipids in sputum of COPD patients and age- and smoking-matched controls without airflow limitation. Since smoking may affect surfactant directly, we included only ex-smokers.

METHODS

Subjects

For the genetic study, we used data from 1390 subjects of the Vlagtwedde/ Vlaardingen cohort who participated in the last survey (1989/1990). This Caucasian general population-based cohort started in 1965 and surveys took place every three years at which participants performed lung function measurements and completed questionnaires. Approximately 25% of the population underwent BHR testing. Cohort details and selection have been described previously [16].

For the sputum lipid analysis, we used material from a previously published cohort of ex-smoking, non-atopic COPD patients and non-atopic healthy ex-smoking controls [17, 18]. Table 1 presents the characteristics of all populations studied. The local university hospital's medical ethics committee approved the study protocols and participants gave written informed consent.

Genoyping

DNA collection and the genotyping protocol are described in the online supplement. *SFTPA1* and *SFTPA2* are homologous genes, and SNP combinations from different alleles: *SFTPA1* 6A, $6A^2$, $6A^3$, $6A^4$ and *SFTPA2* 1A, $1A^0$, $1A^1$, $1A^2$, $1A^3$, $1A^4$, $1A^5$. We determined which *SFTPA1* and *SFTPA2* SNPs tag these alleles and genotyped *SFTPA1* Val19Ala, Leu50Val, Arg219Trp; *SFTPA2* Asn9Thr, Pro91Ala, Ser140Ser. We genotyped functional SNPs in *SFTPB* and *SFTPD*, previously associated with COPD: *SFTPB* Ile131Thr; *SFTPD* Met11Thr, *SFTPD* Thr160Ala [3, 19, 20].

Sputum sample preparation and lipid analysis with high performance liquid chromatography/mass spectrometry

Sputum was collected from 17 COPD patients and 11 ex-smoking healthy controls as described previously [17;18]. Sputum lipid extraction was performed as described by Folch *et al* [21, 22]. Lipid analysis was performed as described previously (see online supplement) [23]. Several lipids were measured including different PC and LPC species, cholesterol and tocopherol.

Table 1: Characteristics of the total Vlagtwedde/Vlaardingen cohort study population, its subpopulations with and without COPD (GOLD stage ≥ II at the last 2 surveys) and the phospholipid study population of COPD patients and controls. Data are presented as median values (interquartile ranges)

	Genetic study			Phospholipid study	
	Vlagtwedde/ Vlaardingen Total population	Vlagtwedde/ Vlaardingen subpopulation without COPD	Vlagtwedde/ Vlaardingen subpopulation with COPD	Ex-smoking COPD patients	Ex-smoking healthy controls
Number of individuals	1390	1204	186	17	11
Age, years	52 (43–60)	50 (42-59)	59 (52 - 64)	63 (57–70)	58 (51–65)
Sex, %male	51	48	74	82	73
Pack-years smoking	9.0 (0–24.0)	6.7 (0-18.0)	21.5 (3.9-32.3)	25.0 (7.5–42.5)	20.0 (10.1–35.0)
FEV_1%pred	93.5(83.8-104.0)	95.8 (87.9-104.5)	71.1 (61.1-77.1)	63.0 (51.0–0.3)	102.0 (93.0–110.0)
FEV_1/VC, %	75.1 (69.6-79.8)	86.6 (72.1-80.5)	60.1 (54.4-65.7)	53.0 (44.5-59.5)	77.0 (72.0-78.0)
ΔFEV_1, ml/yr[†]	-21.1 (-9.9;-32.7)	18.9 (-7.9;-30.0)	-36.8 (-23.5;-50.3)	n.a.	n.a.
FEV_1 values, n	7 (5-8)	7 (5-8)	7 (5-8)	1	1

[†] calculated as last minus first available FEV_1/years participated

Abbreviations: COPD Chronic Obstructive Pulmonary Disease; GOLD Global initiative for Chronic Lung Diseases; FEV_1 Forced Expiratory Volume in one second; VC Vital Capacity

Statistics

Using Linear Mixed Effect (LME) models, we investigate SNP effects on annual decline in Forced Expiratory Volume in 1 second (FEV_1) in the total general population and in subjects with established COPD from this population (defined as GOLD stage II or higher at the last 2 surveys), adjusting for sex, first FEV_1 after age 30, and pack-years smoking [24, 16]. Since smoking affects SP levels, we additionally performed analyses with interaction terms of pack-years and genotypes, and performed stratified analyses by smoking in the total general population. We analyzed genotype effects on first and last available FEV_1 and FEV_1/Vital Capacity (VC) with linear regression (adjusted for sex, age, pack-years, and height). Differences in genotype and haplotype distribution between subjects with and without COPD in the general population were tested using Chi-square tests. BHR and blood eosinophilia were defined dichotomously as respectively $PC_{10} \leq 8$ mg/ml histamine and >275 eosinophils/µl in blood [12]. Genotype and haplotype effects on BHR and eosinophilia were estimated using logistic regression, adjusting for sex, age, smoking and FEV_1 %predicted when appropriate.

We analyzed relations of phospholipid concentrations in sputum with lung function parameters, and blood and sputum eosinophils using Spearman's correlation. Differences in phospholipid concentrations between COPD patients and controls were tested using Mann-Whitney tests.

Statistical analyses were performed using SPSS (version 14.0.1 for Windows), S-Plus (version 7), R (version 1.9.1), and Chaplin [25, 26, 27].

RESULTS

Association of *SFTP* SNPs with lung function parameters, COPD development, BHR and blood eosinophilia in the general population

In the Vlagtwedde/Vlaardingen cohort, minor allelic frequencies for the *SFTP* SNPs were: *SFTPA1* Val19Ala 0.08, Leu50Val 0.10, Arg219Trp 0.09; *SFTPA2* Asn9Thr 0.08, Pro91Ala 0.16, Ser140Ser 0.28; *SFTPB* Ile131Thr 0.47; *SFTPD* Met11Thr 0.45, Thr160Ala 0.60. All SNPs were in Hardy Weinberg equilibrium.

Subjects heterozygous for the *SFTPD* Met11Thr SNP had 2.9 ml/yr excess decline in FEV_1 compared to subjects with the wild type genotype (p=0.034, figure 1). We found significant interaction terms of SNPs *SFTPA1* Leu50Val and *SFTPA1* Ala19Val with pack-years of smoking. Stratified analyses according to smoking showed that SNPs *SFTPA1* Leu50Val and *SFTPD* Met11Thr were significantly associated with FEV_1 decline in ever-smokers, but not in never-smokers (figure 1). In addition, ever-smokers who were heterozygous for SNP *SFTPA1* Ala19Val had 4.8 ml/yr excess FEV_1 decline compared to ever-smokers with the wild type genotype (p=0.046). This effect was not observed in never-smokers. In subjects with established COPD, those who were heterozygous for SNP *SFTPA1* Leu50Val had 11.0 ml/yr less FEV_1 decline compared to those with the wild type genotype (p=0.025, figure 1).

We did not find significant associations between *SFTP* SNPs and BHR (see online supplement, table E1). *SFTPA1* Val19Ala, *SFTPD* Met11Thr and *SFTPD* Thr160Ala, but not the other SNPs, were significantly associated with blood eosinophilia (figure 2). Haplotypes of *SFTPA1* and *SFTPA2* were not associated with BHR or eosinophilia.

Association of surfactant lipids with COPD and lung function parameters

Table 3 shows the surfactant lipids that were detected in the sputum of COPD patients and controls. We found significantly lower levels of PC-C16:0/C16:1, PC-C16:0/C16:0, PC-C16:1/C18:1 or PC-C16:0/C18:2, PC-C16:1/C18:1 and PC-C18:0/C18:2 or PC-C18:1/C18:1 in COPD patients than controls, whereas levels of LPC-C16:0 were significantly higher in COPD patients (table 3).

Higher PC-C16:0/C16:1, PC-C16:1/C18:1 or PC-C16:0/C18:2, PC-C16:0/C18:1 and PC-C18:0/C18:2 or PC-C18:1/C18:1 were significantly associated with higher values of FEV_1/VC and FEV_1 %predicted (table 4). Higher LPC-C16:0 was significantly associated with lower values of FEV_1/VC and FEV_1 %predicted. Higher LPC-C18:1 was associated with lower values of FEV_1/VC (table 4). Figure 3 shows the correlations between PC-C16:0/C16:1 and LPC-C16:0 with FEV_1 %predicted and FEV_1/VC. The other significant correlations of PC and LPC species with lung function parameters resulted in comparable figures. We did not find significant correlations of PCs or LPCs with levels of eosinophils in blood and sputum.

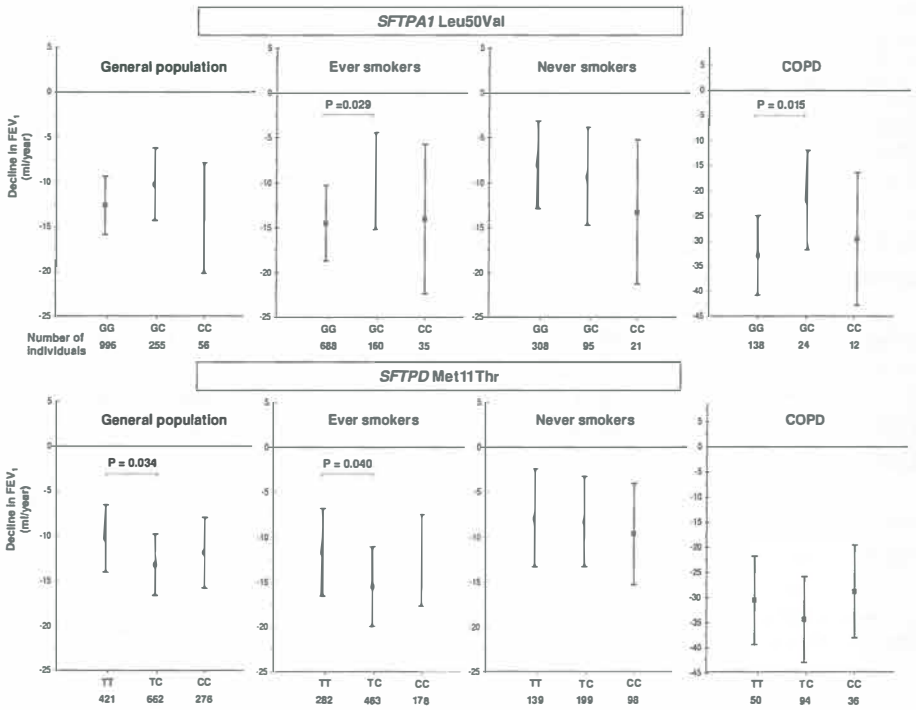


Figure 1: Associations of SFTPA1 Leu50Val and SFTPD Met11Thr with excess FEV_1 decline in the Vlagtwedde/Vlaardingen population, and stratified by ever smokers, never smokers and subjects with COPD (GOLD ≥ II) Mean declines per genotype are adjusted for sex, first FEV_1, age and pack-years of smoking (when appropriate), with 95% confidence intervals.

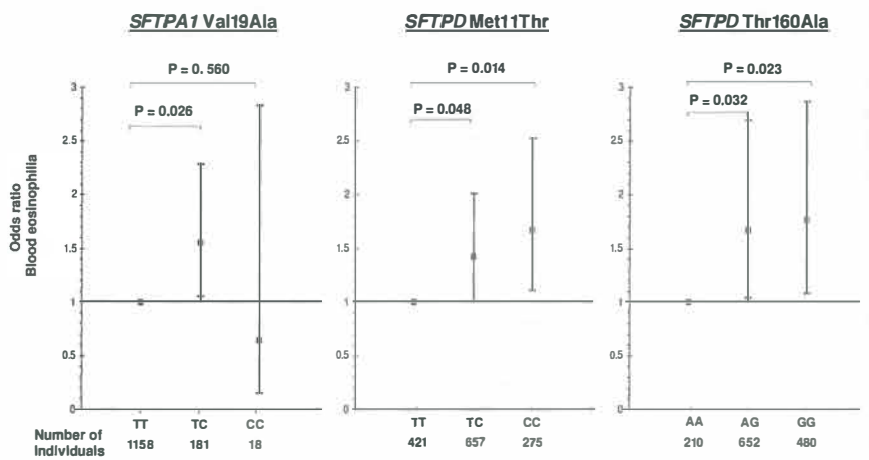


Figure 2: Associations of SFTPA1 Val19Ala, SFTPD Met11Thr, SFTPD Thr160Ala SNPs with blood eosinophilia in the Vlagtwedde/Vlaardingen population Odds ratios with 95% confidence intervals are presented for blood eosinophilia, adjusted for age, sex and pack-years of smoking.

Table 2: Distribution of SNPs in surfactant proteins in subjects with and without COPD (GOLD stage ≥ II) in the Vlagtwedde/Vlaardingen population

SNP		COPD N=186 n (%)	No COPD N=1204 n (%)	P value	SNP		COPD N=186 n (%)	No COPD N=1204 n (%)	P value
SFTPA1	TT	149 (82.6)	970 (85.7)	0.106	*SFTPA2*	CC	90 (50.8)	601 (53.4)	0.816
Val19Ala	TC	31 (17.4)	145 (12.8)		Ser140Ser	CT	71 (40.2)	427 (38.0)	
	CC	0 (0.0)	17 (1.5)			TT	16 (9.0)	97 (8.6)	
SFTPA1	GG	120 (76.0)	743 (75.8)	**0.002**	*SFTPB*	CC	57 (31.1)	313 (27.6)	0.376
Leu50Val	GC	23 (14.7)	203 (20.7)		Ile131Thr	CT	86 (47.6)	594 (52.3)	
	CC	15 (9.3)	34 (3.5)			TT	41 (22.3)	229 (20.1)	
SFTPA1	CC	157 (88.2)	942 (83.3)	0.239	*SFTPD*	TT	41 (22.7)	360 (31.7)	**0.045**
Arg219Trp	CT	20 (11.2)	176 (15.6)		Met11Thr	TC	104 (57.8)	544 (47.9)	
	TT	1 (0.6)	13 (1.1)			CC	35 (19.5)	232 (20.4)	
SFTPA2	AA	65 (34.9)	425 (37.1)	0.732	*SFTPD*	AA	34 (19.2)	171 (15.2)	0.363
Asn9Thr	AC	89 (47.8)	546 (47.7)		Thr160Ala	AG	80 (45.2)	551 (48.8)	
	CC	32 (17.3)	174 (15.2)			GG	63 (35.6)	406 (36.0)	
SFTPA2	GG	131 (72.0)	804 (71.1)	0.590					
Ala91Pro	GC	48 (26.4)	293 (25.9)						
	CC	3 (1.6)	34 (3.0)						

Abbreviations: SNP Single Nucleotide Polymorphism; COPD Chronic Obstructive Pulmonary Disease; GOLD Global initiative for Chronic Lung Diseases; SFTP Surfactant Protein

Table 3: Different lipid concentrations in sputum of COPD patients and controls

MW	Lipid identity	Lipid concentration in sputum (ng/mL)* COPD patients N=17	Controls N=11	P value†
369.8	cholesterol	231,210 (150,872-470,651)	186,323 (51,086-601,240)	0.722
429.9	tocopherol	$18.1*10^6$ ($5.7*10^6$-$31.1*10^6$)	$9.9*10^6$ ($8.3*10^6$-$424.1*10^6$)	0.674
468.8	LPC-C14:0	35.6 (5.9-64.3)	7.1 (4.6-15.8)	0.108
496.8	LPC-C16:0	614.4 (176.2-840.3)	155.4 (123.2-421.8)	**0.010**
522.8	LPC-C18:1	204.6 (61.6-840.3)	55.4 (47.8-187.4)	0.079
707.1	PC-C16:0/C14:0	3,718 (2,905-5,818)	5,242 (3,939-8,203)	0.079
733.2	PC-C16:0/C16:1	4,977 (3,318-13,577)	11,917 (9,547-27,448)	**0.016**
735.2	PC-C16:0/C16:0	12,536 (8,001-18,551)	18,354 (12,682-24,794)	**0.045**
759.2	PC-C16:1/C18:1 or PC16:0/C18:2#	3,077 (2,389-9,166)	9,674 (5,783-15,093)	**0.014**
761.2	PC-C16:1/C18:1	4,736 (3,060-11,340)	11,120 (8,670-18,143)	**0.012**
787.2	PC-C18:0/C18:2 or PC-18:1/C18:1*	1,044 (859-3,465)	3,484 (2,060-5,950)	**0.018**

*Data are presented as median (interquartile range)
†P value from Mann Whitney test for difference in concentration between COPD patients and controls
#based on molecular weight, phospholipid species can be either one of these conformations.
Abbreviations:COPD Chronic Obstructive Pulmonary Disease; MW Molecular Weight; LPC Lysophosphotidylcholine; PC Phosphotidylcholine

Table 4: Significant correlations of lipid species in sputum of COPD patients and controls with lung function parameters

		Positive correlations			Negative correlations		
MW	Lipid identity	parameter	r	P value	parameter	r	P value
369.8	cholesterol	-	-	-	-	-	-
429.9	tocopherol	-	-	-	-	-	-
468.8	LPC-C14:0	-	-	-	FEV_1/VC	-0.516	**0.006**
496.8	LPC-C16:0	-	-	-	FEV_1%pred	-0.452	**0.018**
					FEV_1/VC	-0.646	**<0.001**
522.8	LPC-C18:1	-	-	-	FEV_1/VC	-0.538	**0.004**
707.1	PC-C16:0/C14:0	-	-	-	-	-	-
733.2	PC-C16:0/C16:1	FEV_1%pred	0.424	**0.027**	-	-	-
		FEV_1/VC	0.431	**0.025**	-	-	-
735.2	PC-C16:0/C16:0	-	-	-	-	-	-
759.2	PC-C16:1/C18:1 or	FEV_1%pred	0.458	**0.016**	-	-	-
	PC-16:0/C18:2*	FEV_1/VC	0.434	**0.024**	-	-	-
761.2	PC-C16:1/C18:1	FEV_1%pred	0.478	**0.012**	-	-	-
		FEV_1/VC	0.491	**0.009**	-	-	-
787.2	PC-C18:0/C18:2 or	FEV_1%pred	0.472	**0.013**	-	-	-
	PC-C18:1/C18:1*	FEV_1/VC	0.445	**0.020**	-	-	-

*based on molecular weight, phospholipid species can be either one of these.

Abbreviations:COPD Chronic Obstructive Pulmonary Disease; MW Molecular Weight; r Pearson's correlation coefficient; LPC Lysophosphotidylcholine; PC Phosphotidylcholine; FEV_1 Forced Expiratory Volume in one second; VC Vital Capacity

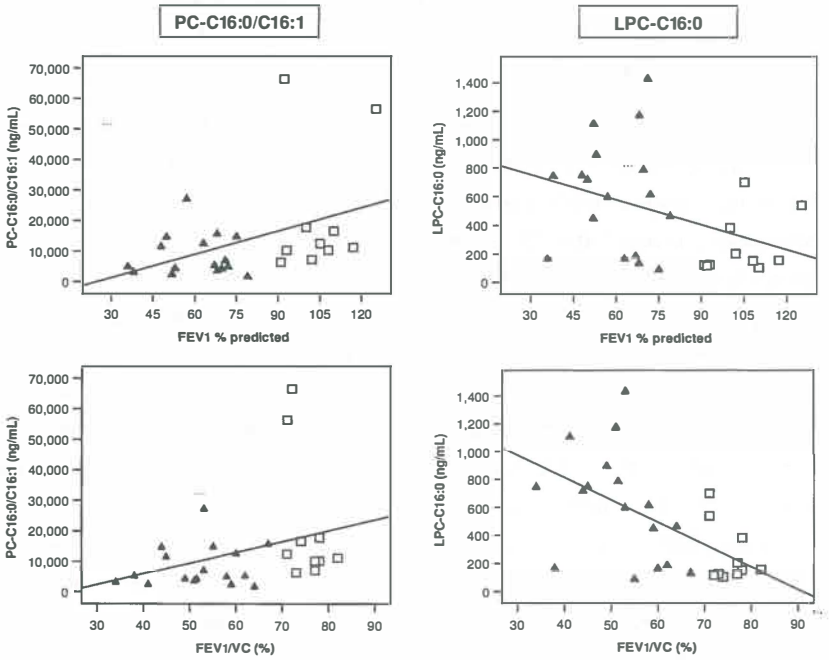


***Figure 3: Correlation of sputum phosphatidylcholine species PC-C16:0/C16:1 and LPC-C16:0 with FEV_1%predicted and FEV_1/VC** Black solid triangles represent COPD patients, open squares controls.*

DISCUSSION

This study sheds light on multiple factors involved in pulmonary surfactant function and composition in relation to COPD. We demonstrate that both genetic factors and lipid variations in surfactant can affect lung function, thereby contributing to the development of COPD. We show for the first time that SNPs in *SFTP* genes are associated with accelerated lung function decline. Moreover, they are associated with high blood eosinophil numbers in the general population, a known risk factor for lower lung function and accelerated lung function decline by itself. In addition, we show that sputum PC fractions are lower in COPD patients than in controls, indicating that degradation of PC is enhanced in COPD. This finding is of clinical importance, since we show that higher levels of degraded PC species are associated with lower lung function.

In our genetic study in the general population, the heterozygous genotype of *SFTPD* Met11Thr was associated with excess FEV_1 decline in the total population and to a lesser extent in ever smokers. We found no significant interaction between the *SFTPD* Met11Thr genotypes and pack-years smoked. Therefore, the detrimental effect of the heterozygous genotype of *SFTPD* Met11Thr SNP on lung function decline seems to be independent of smoking. In contrast, we found a significant positive interaction between the effect of this genotype is modified by smoking. Our stratified analyses according to smoking indeed showed a protective effect of the heterozygous genotype of *SFTPA1* Leu50Val on FEV_1 decline in ever-smokers, whereas this effect was not shown in never-smokers. The heterozygous genotype of *SFTPA1* Leu50Val thus appears to protect the lungs against lung function decline due to smoking.

To our surprise, heterozygosity but not homozygosity for SNPs *SFTPA1* Leu50Val and *SFTPD* Met11Thr was associated with respectively slower and faster FEV_1 decline. Accordingly, the heterozygous genotype of *SFTPA1* Leu50Val was less prevalent and *SFTPD* Met11Thr was more prevalent in subjects with COPD than in those without COPD. This phenomenon is known as molecular heterosis, i.e. the heterozygous effect of a specific genetic polymorphism is significantly greater or smaller than the homozygous or wild-type effect. We cannot speculate about an advantageous role of heterozygosity for *SFTPA1* Leu50Val, since the function of this SNP is unknown. However, we can hypothesize about a detrimental role of *SFTPD* heterozygosity, since it has been shown that the Met11Thr substitution influences the assembly, function and concentration of SP-D [28, 29]. Individuals with Thr/Thr genotypes have significantly lower SP-D serum levels than Met/Met individuals [28, 29]. Met/Met individuals have two distinct forms of SP-D in serum with a high and low molecular weight, whereas Thr/Thr individuals have only the low molecular weight band [29]. The two forms of SP-D display differential binding to intact microorganisms and lipopolysaccharide (LPS). The high molecular weight SP-D may be important in the binding, aggregation and clearance of microorganisms, whereas low

molecular weight SP-D preferentially binds to simpler ligands, like LPS, and may therefore have alternative physiological functions. From this point of view, carrying both allelic variants might be advantageous, since it allows a broader response to microorganisms compared to that of individuals homozygous for either one of the alleles. However, our results point toward the opposite, since the heterozygous genotype was associated with the fastest FEV_1 decline and was more prevalent in COPD subjects. Our data therefore suggest that heterozygous subjects have too low levels of either one of the two molecular weight forms to withstand infections and appropriately clear established infections. This may then result in damage to the lungs and excess decline in lung function.

Surfactant proteins have been previously associated with eosinophil numbers and function, since SP-D regulates eosinophil chemotaxis and degranulation [13]. Eosinophils can produce secretory phospholipases (sPLA2) and eosinophil lysophospholipase, that induce surfactant dysfunction by phospholipid hydrolysis [30]. sPLA2 hydrolyses PC to LPC, which is the substrate for eosinophil lysophospholipases. Interestingly, our study is the first to link the occurrence of higher numbers of blood eosinophils to degradation of phospholipids in both healthy individuals and COPD patients. We show that *SFTPD* SNPs are associated with a high number of eosinophils in peripheral blood. These eosinophils may induce degradation of phospholipids in the lung as indicated by high levels of LPC. Indeed, we found that LPC levels in sputum are higher in COPD patients than in controls. We did not find direct significant associations of *SFTPD* SNPs with levels of sputum LPC, which might be due to the small population size. It would be interesting in further studies to assess such functional relations.

In the lung, surfactant PC is degraded by phospholipase A2 to LPC, which is normally present in small quantities in surfactant. Our COPD patients had significantly lower levels of saturated PC-species and higher levels of LPC-species in sputum than controls, indicating higher degradation of PC in COPD. We observed comparable associations of PC- and LPC-species with lung function parameters. These results are similar to those from a study in asthmatics, showing that higher levels of PC-species were also associated with higher levels of lung function [31]. High levels of LPC in sputum may indicate damage to the surfactant layer, since high concentrations of LPC have been shown to enhance bronchoconstriction in guinea pigs [32, 33]. Our findings furthermore show that high levels of PC are critical in maintaining an optimal lung function as well, because PC levels were higher in our controls than in our COPD patients, and higher PC levels were associated with a better lung function.

Our study has some limitations. For instance, we did not measure levels of surfactant proteins in sputum of the COPD patients and controls, since levels of SP were too low to detect. We therefore can not determine the link between levels of surfactant lipids and proteins in our population of COPD patients and controls. Another limitation of the current study is that our groups of COPD cases and controls were too small to analyze

genetic effects of variants in surfactant proteins on phospholipid content in sputum. Additionally, we did not have sputum of our general population to analyze surfactant fractions and relate them to decline in lung function and specific genotypes.

In conclusion, our study is the first to provide suggestive evidence that alterations in surfactant as caused by effects of both genetic and lipid alterations, affect lung function and development of COPD. We show that genetic variants in *SFTP* genes are associated with higher blood eosinophil numbers, accelerated lung function decline and COPD development in the general population. Furthermore, surfactant phospholipids appear to be important in maintaining an optimal lung function since the PC fraction of lung surfactant is lower and the LPC fraction higher in COPD patients than control subjects. A study by Anzueto *et al.* already showed that treatment with aerosolized surfactant for 2 weeks improves pulmonary function in patients with chronic bronchitis [34]. Since COPD is a disease that is still not appropriately manageable with current available medications, we put forward that therapies directed at improving surfactant levels need attention for benefit of COPD management.

Funding

This study is funded by the Dutch Asthma Foundation, grant 3.2.02.51 and the University of Groningen.

REFERENCES

1. Enhorning G, Duffy LC, Welliver RC. Pulmonary surfactant maintains patency of conducting airways in the rat. *Am J Respir Crit Care Med* 1995;151(2 Pt 1):554-6.

2. Floros J, Thomas NJ, Liu W, *et al*. Family-based association tests suggest linkage between surfactant protein B (SP-B) (and flanking region) and respiratory distress syndrome (RDS): SP-B haplotypes and alleles from SP-B-linked loci are risk factors for RDS. *Pediatr Res* 2006;59(4 Pt 1):616-21.

3. Guo X, Lin HM, Lin Z, *et al*. Surfactant protein gene A, B, and D marker alleles in chronic obstructive pulmonary disease of a Mexican population. *Eur Respir J* 2001;18(3):482-90.

4. Hersh CP, DeMeo DL, Lazarus R, *et al*. Genetic Association Analysis of Functional Impairment in Chronic Obstructive Pulmonary Disease. *Am J Respir Crit Care Med* 2006;173(9):977-84.

5. Lofgren J, Ramet M, Renko M, *et al*. Association between surfactant protein A gene locus and severe respiratory syncytial virus infection in infants. *J Infect Dis* 2002;185(3):283-9.

6. Marttila R, Haataja R, Guttentag S, *et al*. Surfactant protein A and B genetic variants in respiratory distress syndrome in singletons and twins. *Am J Respir Crit Care Med* 2003;168(10):1216-22.

7. Tashkin DP, Altose MD, Bleecker ER, *et al*. The lung health study: airway responsiveness to inhaled methacholine in smokers with mild to moderate airflow limitation. The Lung Health Study Research Group. *Am Rev Respir Dis* 1992;145(2 Pt 1):301-10.

8. Jansen DF, Rijcken B, Schouten JP, *et al*. The relationship of skin test positivity, high serum total IgE levels, and peripheral blood eosinophilia to symptomatic and asymptomatic airway hyperresponsiveness. *Am J Respir Crit Care Med* 1999;159(3):924-31.

9. Balzano G, Stefanelli F, Iorio C, *et al*. Eosinophilic inflammation in stable chronic obstructive pulmonary disease. Relationship with neutrophils and airway function. *Am J Respir Crit Care Med* 1999;160(5 Pt 1):1486-92.

10. Wang X, Mensinga TT, Schouten JP, *et al*. Determinants of maximally attained level of pulmonary function. *Am J Respir Crit Care Med* 2004;169(8):941-9.

11. Lebowitz MD, Postma DS, Burrows B. Adverse effects of eosinophilia and smoking on the natural history of newly diagnosed chronic bronchitis. *Chest* 1995;108(1):55-61.

12. Mensinga TT, Schouten JP, Weiss ST, *et al*. Relationship of skin test reactivity and eosinophilia to level of pulmonary function in a community-based population study. *Am Rev Respir Dis* 1992;146(3):638-43.

13. von Bredow C, Hartl D, Schmid K, *et al*. Surfactant protein D regulates chemotaxis and degranulation of human eosinophils. *Clin Exp Allergy* 2006;36(12):1566-74.

14. Honda Y, Takahashi H, Kuroki Y, *et al*. Decreased contents of surfactant proteins A and D in BAL fluids of healthy smokers. *Chest* 1996;109(4):1006-9.

15. Wang Y, Krull IS, Liu C, *et al*. Derivatization of phospholipids. *J Chromatogr B Analyt Technol Biomed Life Sci* 2003;793(1):3-14.

16. van Diemen CC, Postma DS, Vonk JM, *et al*. A disintegrin and metalloprotease 33 polymorphisms and lung function decline in the general population. *Am J Respir Crit Care Med* 2005;172(3):329-33.

17. Rutgers SR, Timens W, Kaufmann HF, *et al*. Comparison of induced sputum with bronchial wash, bronchoalveolar lavage and bronchial biopsies in COPD. *Eur Respir J* 2000;15(1):109-15.

18. Rutgers SR, Postma DS, ten Hacken NH, *et al*. Ongoing airway inflammation in patients with COPD who do not currently smoke. *Thorax* 2000;55(1):12-8.

19. Hersh CP, DeMeo DL, Lange C, *et al*. Attempted Replication of Reported Chronic Obstructive Pulmonary Disease Candidate Gene Associations. *Am J Respir Cell Mol Biol* 2005;33(1):71-8.

20. Seifart C, Plagens A, Brodje D, *et al*. Surfactant protein B intron 4 variation in German patients with COPD and acute respiratory failure. *Dis Markers* 2002;18(3):129-36.

21. Folch J, Lees M, Sloane SGH. A simple method for the isolation and purification of total lipides from animal tissues. *J Biol Chem* 1957;226(1):497-509.

22. von BC, Birrer P, Griese M. Surfactant protein A and other bronchoalveolar lavage fluid proteins are altered in cystic fibrosis. *Eur Respir J* 2001;17(4):716-22.

23. Barroso B, Bischoff R. LC-MS analysis of phospholipids and lysophospholipids in human bronchoalveolar lavage fluid. *J Chromatogr B Analyt Technol Biomed Life Sci* 2005;814(1):21-8.

24. From the Global Strategy for the Diagnosis, Management and Prevention of COPD, Global Initiative for Chronic Obstructive Lung Disease (GOLD) 2006. Available from: http://www.goldcopd.org.

25. R development Core Team. R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing; 2004.

26. Case-Control Haplotype Interference (CHAPLIN), version 1.2,. 2006.

27. Epstein MP, Satten GA. Inference on haplotype effects in case-control studies using unphased genotype data. *Am J Hum Genet* 2003;73(6):1316-29.

28. Heidinger K, Konig IR, Bohnert A, *et al*. Polymorphisms in the human surfactant protein-D (SFTPD) gene: strong evidence that serum levels of surfactant protein-D (SP-D) are genetically influenced. *Immunogenetics* 2005;57(1-2):.(1-2):1-7.

29. Leth-Larsen R, Garred P, Jensenius H, *et al*. A common polymorphism in the SFTPD gene influences assembly, function, and concentration of surfactant protein D. *J Immunol* 2005;174(3):1532-8.

30. Kwatia MA, Doyle CB, Cho W, *et al*. Combined activities of secretory phospholipases and eosinophil lysophospholipases induce pulmonary surfactant dysfunction by phospholipid hydrolysis. J Allergy Clin Immunol 2007;119(4):838-47.

31. Wright SM, Hockey PM, Enhorning G, *et al*. Altered airway surfactant phospholipid composition and reduced lung function in asthma. *J Appl Physiol* 2000;89(4):1283-92.

32. Mehta D, Gupta S, Gaur SN, *et al*. Increased leukocyte phospholipase A2 activity and plasma lysophosphatidylcholine levels in asthma and rhinitis and their relationship to airway sensitivity to histamine. *Am Rev Respir Dis* 1990;142(1):157-61.

33. Nobata K, Kurashima K, Fujimura M, *et al*. Inhaled lysophosphatidylcholine provokes bronchoconstriction in guinea pigs in vivo. *Eur J Pharmacol* 2005;520(1-3):150-5.

34. Anzueto A, Jubran A, Ohar JA, *et al*. Effects of aerosolized surfactant in patients with stable chronic bronchitis: a prospective randomized controlled trial. *JAMA* 1997;278(17):1426-31.

Online Data Supplement to paper:

Pulmonary surfactant: genetic and lipidomic risk factors for Chronic Obstructive Pulmonary Disease

Cleo C van Diemen, Dirkje S Postma, Begona Barroso, Judith M Vonk, Rainer Bischoff, H Marike Boezen

METHODS

Sputum sample preparation and analysis of lipids in sputum with high performance liquid chromatography (HPLC)/mass spectrometry (MS)

Sputum was collected from 16 COPD patients and 11 ex-smoking healthy controls as described previously. [1, 2] Extraction of lipids from sputum was performed by the method described by Folch *et al.* [3, 4]. In short, 1.5 ml of a mixture of chloroform/methanol/50 mM acetic acid 8/4/0.5 (v/v/v) was mixed with 0.5 ml of sputum supernatant and manually shaken for 2 min. The mixture was centrifuged for 5 min at 1500 g, and the organic layer separated. The aqueous phase was extracted again and both organic extracts were pooled together and evaporated to dryness under a stream of nitrogen. The residue was reconstituted in 50 µl of DMSO by ultrasonication and vortex mixing. 5 µl were injected in the HPLC/MS system. Glassware was used throughout the whole sample preparation procedure.

The HPLC part of the analytical system consisted of an Agilent Series 1100 capillary LC system (Waldbronn, Germany) comprising a degasser, a binary pump, a thermostatted autosampler and a thermostatted column compartment. Chromatographic separation took place on a reversed-phase column (Poroshell 300SB-C8, 5 µm, 1 mm i.d. × 75 mm length) (Agilent, Waldbronn, Germany). Mobile phase A consisted of 0.1% formic acid in ultrapure water. Mobile phase B was 0.1% formic acid in acetonitrile. Separation was performed with an increasing gradient of B (25–95% in 50 min). Flow rate employed was 60 µl/min and the column temperature was 60 °C. The analytes were detected by an Agilent SL ion-trap mass spectrometer equipped with an electrospray ionization (ESI) source operated in the positive mode. MS data were acquired over a scan range of 100–1200 amu and 5500 m/z per second scan rate. Several lipids were measured including different phosphatidylcholine (PC) and lysophosphatidylcholine (LPC) species, cholesterol and tocopherol.

DNA collection

DNA in the Vlagtwedde/Vlaardingen cohort was extracted from neutrophil depot as previously published. [5] We used the classical chloroform extraction for genomic DNA isolation from the sputum samples and checked the DNA for purity and concentration with the NanoDrop® ND-1000 UV-Vis Spectrophotometer (Nanodrop Technologies, Wilmington, Delaware, USA).

Genotyping

Primers and probes were obtained from Applied Biosystems TaqMan® SNP Genotyping Assays (Nieuwekerk aan de IJssel, The Netherlands), using the Assay-on demand service. Sequences of primers and probes are available upon request. Reactions were performed in 5 µl volumes and contained 10 ng DNA, 1x Taqman Universal Mastermix (Applied Biosystems), 200 nM of each probe and 900 nM of each primer. Cycling conditions on the ABI PRISM® 7900HT Sequence Detection System (Applied Biosystems) were 10 min 95°C followed by 40 cycles of 15 seconds 95°C and 1 minute 60°C. End-point fluorescence was measured immediately after cycling. Alleles were assigned using SDS 2.1 software (Applied Biosystems). We regenotyped 6% of the samples and found no errors in the genotypes.

REFERENCES

1. Rutgers SR, Timens W, Kaufmann HF, *et al*. Comparison of induced sputum with bronchial wash, bronchoalveolar lavage and bronchial biopsies in COPD. *Eur Respir J* 2000;15(1):109-15.
2. Rutgers SR, Postma DS, ten Hacken NH, *et al*. Ongoing airway inflammation in patients with COPD who do not currently smoke. *Thorax* 2000;55(1):12-8.
3. Folch J, Lees M, Sloane SGH. A simple method for the isolation and purification of total lipides from animal tissues. *J Biol Chem* 1957;226(1):497-509.
4. von BC, Birrer P, Griese M. Surfactant protein A and other bronchoalveolar lavage fluid proteins are altered in cystic fibrosis. *Eur Respir J* 2001;17(4):716-22.
5. van Diemen CC, Postma DS, Vonk JM, *et al*. A disintegrin and metalloprotease 33 polymorphisms and lung function decline in the general population. *Am J Respir Crit Care Med* 2005;172(3):329-33.

Table E1: Association of *SFTP* SNPs with blood eosinophilia and BHR in the total Vlagtwedde/Vlaardingen population

SNP		Eos OR*	95% C.I.	P value	BHR PC10≤ 8 mg/ml OR†	95% C.I.	P value
SFTPA1	TC	**1.55**	**1.05 - 2.28**	**0.026**	1.07	0.58 - 2.00	0.818
Val19Ala	CC	0.64	0.15 - 2.83	0.560	0.81	0.12 - 0.95	0.836
SFTPA1	GC	1.23	0.86 - 1.75	0.253	1.55	0.87- 2.76	0.134
Leu50Val	CC	0.80	0.37 - 1.74	0.579	1.13	0.44 - 2.90	0.807
SFTPA1	CT	0.64	0.41 - 1.01	0.057	0.77	0.41 - 1.45	0.420
Arg219Trp	TT	1.57	0.49 - 4.96	0.442	-	-	-
SFTPA2	AC	0.90	0.66 - 1.24	0.531	0.79	0.48 - 1.30	0.359
Asn9Thr	CC	1.12	0.74 - 0.70	0.579	1.04	0.52 - 2.07	0.913
SFTPA2	GC	1.22	0.88 - 1.67	0.225	1.00	0.61 - 1.65	0.313
Ala91Pro	CC	0.54	0.19 - 1.55	0.255	0.41	0.07 - 2.32	0.991
SFTPA2	CT	0.91	0.67 - 1.22	0.514	1.17	0.72 - 1.89	0.524
Ser140Ser	TT	0.66	0.37 - 1.16	0.149	1.63	0.73 - 3.63	0.230
SFTPB	CT	1.06	0.76 - 1.49	0.733	1.75	0.45 - 1.27	0.287
Ile131Thr	TT	1.08	0.72 - 1.64	0.703	0.73	0.38 - 1.39	0.343
SFTPD	TC	**1.42**	**1.00 - 2.01**	**0.048**	1.05	0.63 - 1.77	0.844
Met11Thr	CC	**1.67**	**1.11 - 2.52**	**0.014**	1.75	0.94 - 3.23	0.077
SFTPD	AG	**1.67**	**1.04 - 2.69**	**0.032**	1.56	0.77 - 3.17	0.215
Thr160Ala	GG	**1.76**	**1.08 - 2.86**	**0.023**	1.89	0.92 - 3.89	0.085

*Odds ratio compared to wild-type genotype, adjusted for sex, age, pack-years of smoking.

† Odds ratio compared to wild-type genotype, adjusted for sex, age, pack-years of smoking, FEV_1 percent predicted.

Abbreviations: SFTP Surfactant Protein; SNP Single Nucleotide Polymorphism; BHR Bronchial Hyperresponsiveness; Eos Eosinophilia; OR Odds Ratio; CI Confidence Interval

CHAPTER 11

Novel strategy to identify genetic risk factors for COPD severity: genetically isolated populations

Cleo C van Diemen
Dirkje S Postma
Yurii S Aulchenko
Peter JLM Snijders
Ben A Oostra
Cornelia M van Duijn
H Marike Boezen

Submitted for publication

ABSTRACT

Background

Genetic determinants for Chronic Obstructive Pulmonary Disease (COPD) are difficult to study, since COPD develops at later ages and involves multiple genetic and environmental factors. Studies using genetically isolated populations with limited genetic variation may be useful in COPD genetics, but are thus far lacking.

Aim

To study associations between Single Nucleotide Polymorphisms (SNPs) in candidate genes and severity of airway obstruction in COPD in a genetically isolated population, and to replicate results in the general population.

Methods

We collected spirometry and questionnaire information from 157 subjects with a doctor's diagnosed COPD from the Genetic Research in Isolated Population (GRIP) study. We analyzed 32 SNPs in 13 candidate genes: *ADAM33*, *TGFβ1*, *MMP1*, *MMP2*, *MMP9*, *MMP12*, *TIMP1*, *SFTPA1*, *SFTPA2*, *SFTPB*, *SFTPD*, *GSTP1* and *HMOX1*. We analyzed FEV_1, FVC, and FEV_1/FVC levels in COPD patients using restricted maximum likelihood linear mixed modeling, accounting for pedigree structure. We replicated significant associations in 351 COPD patients selected from the general Vlagtwedde/Vlaardingen study.

Results

We studied 106 spirometry confirmed COPD GOLD stage ≥ I subjects, members of an extended pedigree including 6,175 people. Seven SNPs in *ADAM33*, *TGFB1*, *SFTPA1*, *SFTPA2* and *SFTPD* were significantly associated with FEV_1/FVC level. The *ADAM33* and *TGFB1* associations with FEV_1/FVC were replicated in the Vlagtwedde/Vlaardingen population, with similar effect sizes. Allele frequencies were similar in both populations.

Discussion

We show for the first time that a genetic isolate can be used to determine genetic risk factors for lung function level, which can be replicated in COPD patients from an independent population.

INTRODUCTION

Chronic Obstructive Pulmonary Disease (COPD) is the third cause of death worldwide and is expected to increase in prevalence in the forthcoming decades [1;2]. The disease has a large personal, societal, and economic impact. COPD is characterized by chronic airway inflammation, airway remodeling, and airflow limitation that is not fully reversible. Since not all smokers develop COPD, genetic susceptibility has to/may play a role in development of this disease, in addition to environmental factors. The genetic determinants for COPD are difficult to study, since COPD is a disease that becomes clinically manifest only at later ages, when parents of COPD patients have already died, and their children are likely to young to manifest airway obstruction. This limits the option to perform family based genetic research. Moreover, published studies frequently use various definitions of disease status, which makes it difficult to compare their results. Therefore, it makes sense to choose a robust phenotype to define COPD like the level of lung function, which can be more easily compared between studies. Moreover, a low level of lung function is a predictor for mortality from COPD [3-5].

Another complicating factor in studies on genetics of COPD is that COPD is considered a complex genetic trait, i.e. multiple, possibly interacting, genetic and environmental factors are involved. Therefore it has advantages to try and identify risk genes in populations that are relatively genetically and environmentally homogeneous, such as genetically isolated populations. Due to the small number of founders and drift in genetically isolated populations, the genetic variation is reduced [6]. However, these processes raise the question whether findings can be extrapolated to the general population. Previous simulation studies suggest that this is the case for common variants with a frequency of >1% [6], but no empirical evidence is available.

We conducted a candidate gene study for level of airflow limitation in patients with COPD who were ascertained as part of the Genetic Research in Isolated Populations (GRIP) study that is conducted in a young genetically isolated population from the southwestern part of the Netherlands. All patients were genotyped using 32 Single Nucleotide Polymorphisms (SNPs) in 13 candidate genes for COPD, chosen based on their previously published association with either COPD, level of lung function, or lung function decline as reported in the open population. Extensive genealogy information was collected resulting in an extremely large and complex pedigree of 6,175 members. Finally, we studied 1390 individuals from the general Dutch population, including 351 patients with COPD, to establish whether our findings could be replicated in the general population. In both studies, we investigated whether the severity of the disease, as reflected by lung function reduction is genetically influenced in established COPD.

METHODS

Study populations

Our study is part of the GRIP program [7;8]. GRIP is based in a recent genetically isolated population from the southwestern of the Netherlands, which was founded in the middle of the 18th century by approximately 150 individuals and was genetically isolated until the last few decades. The population now includes approximately 20,000 inhabitants in 8 adjacent communities. GRIP participants are generally related via multiple lines of descent and are inbred via multiple consanguineous loops [9;10].

We invited subjects with general practitioner's diagnosed COPD to the research center to undergo spirometry and complete a questionnaire [11]. Spirometry was performed by trained pulmonary research technicians using a Pneumotagograph (Viasys, formerly Jaeger Spirometer system). Predicted values for FEV_1 were calculated using adjusted Quanjer-equations for Caucasian subjects [12]. We isolated DNA from blood using Puregene® DNA Purification Kits (Gentra, Inc, Minneapolis, USA). All participants gave written informed consent.

To test the findings from GRIP in the general population, we used cross-sectional data from the general population-based Vlagtwedde/Vlaardingen cohort. Questionnaires, spirometry and DNA were collected [13;14]. For this study, we selected 351 subjects according to the Global Initiative for Chronic Obstructive Lung Disease (GOLD) criteria with GOLD stage ≥ I COPD at the last 1989/1990 survey (table 1) [18]. Since we did not collect FVC values in the Vlagtwedde/Vlaardingen cohort, we used VC values for the analyses.

Genotyping

We have genotyped SNPs in candidate genes for lung function and COPD, based on their previously published significant associations (table 2). Genotyping was performed using Applied Biosystems TaqMan® SNP Genotyping Assays (Nieuwekerk aan de IJssel, The Netherlands). Sequences of primers and probes are available on request.

Statistical analysis

To analyze pedigree data, we used Measured Genotype (MG) approach [15], which models quantitative traits as

$$y_i = \mu + kg_i + \sum_j \beta_j c_{ji} + G_i + e_i$$

where y_i: the phenotype of the i-th individual, g: the vector of genotypes at the marker under study, k: is the marker genotype effect, c_{ij}: the value of the j-th covariate or fixed effect for the individual i, β_j: an estimate of the j-th fixed effect or covariate, and G_i and e_i are random additive polygenic and residual effects, respectively. The random effects are assumed to follow multivariate normal distribution with mean zero. The variance for the

polygenic effects is defined as $\Phi\sigma_G^2$, where Φ is the relationship matrix and σ_G^2 the additive genetic variance due to polygenes. For the residual random effects, the variance is defined as $I\sigma_e^2$, where I is the identity matrix and σ_e^2 the residual variance. Because the pedigree under analyses was very large, we used fast GRAMMAR approximation to the full MG approach [16]. The GRAMMAR consist of a fast though conservative test at screening stage, followed up with full MG analysis of polymorphisms which pass the relaxed (P<0.1) screening significance threshold. All analysis involving pedigree were performed using ASReml v 2.0 [17]– a package for linear mixed model analysis using restricted maximum likelihood.

In the Vlagtwedde/Vlaardingen population, we tested significant associations using linear regression analyses. All analyses were adjusted for age, height and sex.

Table 1: Differences in characteristics of the GRIP and Vlagtwedde study populations
Data are presented as median (range).

	GRIP ≥GOLD I N=106	Vla/Vla ≥GOLD I N=351	P value *	GRIP ≥GOLD II N=70	Vla/Vla ≥GOLD II N=167	P value *
Age	65.5 (37-84)	58.0 (35-76)	<0.001	66.0 (43-82)	59.0 (35-76)	<0.001
Sex m/f, n	53/53	244/107	<0.001	37/30	122/45	0.002
Smoking,						
% Never	3.8	18.8	<0.001	2.9	16.2	0.020
% Ex	34.0	35.9		38.3	33.5	
% Current	58.5	45.3		58.8	50.3	
Pack-years	32.3 (0-120)	21.4 (0-262)	0.002	39.0 (0-120)	26.0 (0-262)	0.099
FEV_1 % of predicted	71.2 (26.4-132.8)	80.7 (36.0-115.0)	0.001	63.9 (26.4-79.0)	69.9 (36.0-79.8)	0.003
FEV_1/FVC #	57.7 (27.7-69.2)	64.9 (29.0-69.9)	<0.001	53.1 (27.7-68.9)	59.2 (29.4-69.8)	<0.001
Chronic cough, %	60.6	14.5	<0.001	60.9	22.2	<0.001
Chronic phlegm, %	50.0	10.5	<0.001	50.7	15.0	<0.001

*P value for difference between study populations derived from X^2-test for comparison of discrete variables; Mann-Whitney U test for test between continuous variables
#In GRIP FEV_1/FVC was used; in Vlagtwedde/Vlaardingen FEV_1/VC
Abbreviations: GRIP Genetic Research in Isolated populations; Vla/Vla Vlagtwedde/Vlaardingen; GOLD Global initiative for Obstructive Lung Diseases; m male; f female; FEV_1 Forced Expiratory Volume in 1 second; FVC Forced Vital Capacity

Table 2: Candidate genes and single nucleotide polymorphisms genotyped in the study population

Gene	Description of gene	SNPs genotyped rs numbers	alternative name
ADAM33	A Disintegrin And Metalloprotease 33: exact function unknown; identified by genome wide screen as susceptibility gene for asthma; associated with decline in FEV_1 and development of COPD in the general population	rs17548913 rs17548907 rs3918396 rs528557 rs597980 rs2280091 rs2280090 rs2787094	ADAM33F+1 ADAM33Q-1 ADAM33S1 ADAM33S2 ADAM33ST+5 ADAM33T1 ADAM33T2 ADAM33V4
TGFβ1	Transforming Growth Factor-β1:a chemotactic cytokine for fibroblasts, inducing synthesis of matrix proteins and glycoproteins and inhibiting collagen degradation by induction of protease inhibitors and reduction of metalloproteases; TGF-β1 levels are increased in COPD; SNPs have been associated with COPD	rs1800469 rs1982073 rs6957	TGFB1C-509T TGFB1Leu10Pro TGFB13UTR
SFTPA1	Surfactant protein A1: Surfactant proteins (SP) are involved in the first response to microorganisms in the lung, regulation of inflammation and structure of alveoli. SP reduce surface tension at the air–liquid interface and therefore prevent alveolar collapse during expiration	rs1059047 rs1136450 rs4253527	SPA1Val19Ala SPA1Leu50Val SPA1Arg219Trp
SFTPA2	Surfactant protein A2: idem SP-A1, homologous gene	rs1059046 rs17886395 rs1965707	SPA2Asn9Thr, SPA2Pro91Ala, SPA2Ser140Ser
SFTPB	Surfactant protein B: hydrophobic component of pulmonary surfactant	rs1130866	SPBIle131Thr
SFTPD	Surfactant protein D: a C-type lectin present in pulmonary surfactant and several other mucosal surfaces. It modulates innate immunity, allergic response, expression of matrix metalloproteases, alveolar wall remodeling, emphysema, fibrosis and lipid and macrophage homeostasis	rs721917 rs2243639	SPDMet11Thr SPDThr160Ala
MMP1	Matrix Metalloprotease 1: an interstitial collagenase involved in tissue remodeling and repair associated with lung development and inflammation. Levels are increased in sputum of COPD patients compared to healthy controls.	rs1799750	MMP1G-1607GG
MMP2	Matrix Metalloprotease 2: a type IV collagenase specifically cleaving type IV collagen, the major structural component of basement membranes	rs243865	MMP2C-1306T
MMP9	Matrix Metalloprotease 9: a gelatinase B involved in tissue remodeling; smokers with airway obstruction show higher MMP9 expression than smokers without COPD and non-smokers	rs3918278 rs6065912 rs8113877	mmp9_rs3918278 mmp9_rs6065912 mmp9_rs8113877
MMP12	Matrix Metalloprotease 12: a human macrophage elastase involved in degradation of extracellular matrix in lungs of patients with COPD	rs2276109 rs652438	MMP12A-82G MMP12Asn357Ser
TIMP1	Tissue Inhibitor of Matrix Metalloprotease 1: inhibitor of several MMPs, including MMP1, MMP9 and MMP12. X-chromosomal	rs11551797 rs4898	timp1Ile158 timp1Phe124
HMOX1	Hemoxygenase 1: role in oxidant-antioxidant balance in the lung	rs2071747	HO1Asp7His
GSTP1	Glutathione S-transferase P1: role in oxidant-antioxidant balance in the lung	rs1695 rs1138272	gstp1Ile105Val gstp1Ala114Val

RESULTS

GRIP study population

We ascertained 157 individuals who were diagnosed with COPD by their general practitioners. Spirometry measures confirmed COPD in 106 subjects, i.e. 36 with COPD GOLD stage I (FEV_1/FVC<70%), 43 with GOLD stage II (FEV_1/FVC <70% and $FEV_1$50-80% predicted) and 37 with GOLD ≥ III (FEV_1/FVC<70% and FEV_1/FVC<50% predicted) [18]. Fifty-one subjects could not be defined as having COPD according to their spirometry (see online supplement table E1) and were therefore excluded from the analyses. The characteristics of the GRIP COPD population are shown in table 1. We determined the familial relationship of the 106 subjects with COPD in the larger GRIP study database. This resulted in a large extended pedigree structure of 6,175 members.

Association of genes with lung function parameters

We analyzed the effects of SNPs in the studied genes on FEV_1 % predicted, FVC and FEV_1/FVC in the 106 subjects with COPD defined as GOLD stage ≥ I. When analyzing FEV_1/FVC, 7 SNPs located in *ADAM33*, *TGFβ1*, *SFTPA1*, *SFTPA2* and *SFTPD* passed the screening stage and were followed up using full MG model. This analysis revealed significant association between FEV_1/FVC and 7 SNPs in genes *ADAM33*, *TGFβ1*, *SFTPA1*, *SFTPA2* and *SFTPD* (table 3).

None of the SNPs was associated with FEV_1 % predicted or FVC.

Replication of GRIP findings in the Vlagtwedde/Vlaardingen population

We found no significant associations of any of the SNPs with FEV_1/VC or with FEV_1% predicted within subjects with GOLD stage ≥ I in Vlagtwedde/Vlaardingen. The GRIP COPD GOLD stage ≥ I patients were significantly different compared to the Vlagtwedde/Vlaardingen COPD GOLD stage ≥ I patients for amongst others age, lung function, and smoking history COPD (see table 1). The GRIP COPD patients had more severe COPD than the Vlagtwedde/Vlaardingen GOLD stage ≥ I patients. To compare COPD populations with more similar phenotypes, we performed additional linear regression analyses in the GRIP and the Vlagtwedde/Vlaardingen populations specifically in subjects with GOLD stage ≥ II (characteristics, see table 1). We found that in both populations SNPs *ADAM33* ST+5, *TGFβ1* C-509T and *TGFβ1* Leu10Pro were significantly or suggestively associated with FEV_1/(F)VC ($p<0.10$, table 3).

Table 3: Significant associations of SNPs with FEV_1/FVC in subjects from the GRIP study with COPD GOLD stage ≥ I, and subjects from the Vlagtwedde/Vlaardingen population with COPD GOLD stage ≥ II Estimates are adjusted for sex, height, age and pedigree structure (applicable in GRIP)

SNP	Comparison *	N	Estimate GRIP GOLD ≥ I	s.e.	P value	N	linear regression estimate Vla/Vla GOLD ≥ II	s.e.	P value
ADAM33	Wt	15	ref	-	-	27	ref	-	-
ST+5	het vs. wt	53	5.2	2.8	**0.070**	86	2.9	1.7	**0.082**
	hom vs. wt	36	5.4	3.0	**0.064**	51	0.8	1.8	0.668
TGFβ1	Wt	46	ref	-	-	94	ref	-	-
C-509T	het vs. wt	45	-3.7	1.9	**0.055**	60	-1.3	1.2	0.298
	hom vs. wt	12	-8.0	3.0	**0.008**	8	-5.0	2.8	**0.070**
TGFβ1	Wt	35	ref	-	-	68	ref	-	-
Leu10Pro	het vs. wt	50	-3.6	2.1	**0.087**	65	-0.8	1.3	0.952
	hom vs. wt	14	-7.1	3.0	**0.019**	17	-4.5	2.0	**0.028**
SFTPA1	Wt	70	ref	-	-	12	ref	-	-
Leu50Val	het vs. wt	20	-5.1	2.4	**0.033**	20	2.7	1.9	0.159
	hom vs. wt	6	6.9	4.1	**0.093**	11	1.8	2.4	0.474
SFTPA2	Wt	70	ref	-	-	11	ref	-	-
Pro91Ala	het vs. wt	32	-1.1	2.1	0.589	41	-1.1	1.4	0.423
	hom vs. wt	3	-12.5	5.6	**0.026**	3	0.4	4.5	0.923
SFTPD	Wt	40	ref	-	-	44	ref	-	-
Met11Thr	het vs. wt	42	-4.2	2.1	**0.048**	85	-0.9	1.4	0.512
	hom vs. wt	20	-5.8	2.7	**0.031**	31	-0.3	1.8	0.888
SFTPD	Wt	37	ref	-	-	54	ref	-	-
Ala160Thr	het vs. wt	47	4.3	2.1	**0.041**	73	2.1	1.4	0.112
	hom vs. wt	16	6.1	3.0	**0.042**	29	-1.6	1.8	0.376

* Model of inheritance: general model: het vs. wt: effect of heterozygous genotype compared to wild type genotype; hom vs. wt: effect of homozygous mutant genotype compared to wild type genotype

Abbreviations: GRIP Genetic Research in Isolated populations; Vla/Vla Vlagtwedde/Vlaardingen; GOLD Global initiative for Obstructive Lung Diseases;ADAM33 A Disintegrin and Metalloprotease 33; TGFβ1 Transforming Growth Factor β1; SFTP Surfactant Protein; ref reference; se standard error

Discussion

Our study is the first to use a genetically isolated population to analyze genetic effects on level of lung function in COPD. Interestingly, we found significant effects of SNPs in COPD candidate genes on the severity of COPD, assessed by lung function in subjects with COPD even though our study population is small. Our results show that levels of FEV_1/FVC, a measure of airway obstruction, are genetically influenced in established COPD. This means that even within patients with phenotypical COPD, we can identify genotypes that are associated with severity of the disease. This is of clinical importance since low lung function levels have been shown to predict mortality of COPD not only in the general population, but also within COPD patients [3-5].

The SNPs that were associated with level of FEV_1/FVC in the current study have been shown to affect severity of COPD in previous studies. The wild type allele of the *ADAM33* ST+5 SNP has been associated with more severe bronchial hyperresponsiveness and more inflammation in bronchial biopsies and sputum in another Dutch cohort of patients with mild to moderate COPD [19]. In patients with severe COPD, the *ADAM33* ST+5 wild type allele has been associated with excess decline in FEV_1 [20]. The current finding that this genotype is associated with lower FEV_1/FVC levels in patients with COPD complements the previous studies, and suggests that the *ADAM33* ST+5 SNP is a modifier SNP for COPD severity.

The *TGFβ1* SNPs that were associated with FEV_1/FVC in our populations have previously been associated with development of COPD or with lower levels of FEV_1 and FEV_1/FVC in several [21-23], but not all previous studies [14;24;25]. Our results (in both the genetically isolated and general population) thus confirm the former studies that implicate a role of *TGFβ1 i*n the severity of airflow limitation. The associated *SFTPA1, SFTPA2* and *SFTPD* SNPs have been associated with COPD previously [26;27]. We now for the first time show that these SNPs may also play a role in severity of COPD. This is plausible, since surfactant proteins decrease surface tension at the air–liquid interface and, therefore, reduce the tendency of alveoli to collapse during expiration. The latter contributes to the severity of airway obstruction as measured by FEV_1/FVC.

We found no significant associations of *MMP1, MMP2, MMP9, MMP12, TIMP1, SFTPB, GSTP1* and *HMOX1* with level of lung function in COPD patients. This does not imply that these genes do not play a role in COPD whatsoever. So far, no studies have analyzed genetic effects on the severity of airway obstruction *within* patients with established COPD. Our study shows that SNPs in *ADAM33, TGFβ1, SFTPA*1, *SFTPA2* and *SFTPD* may be important in progression of COPD, whereas the SNPs in other genes, i.e. *MMP1, MMP2, MMP9, MMP12, TIMP1, GSTP1* and *HMOX1,* may simply constitute genes that are important in development of COPD.

One important advantage of testing genes in a genetically isolated population is that it provides the opportunity to find genes associated with disease in a relatively small sample size due to increased homogeneity of the population. Thus for a lower cost and effort, one can test many genes as to their significance in contributing to disease severity, which subsequently can then be replicated in a larger sample of the general population. The most important requirement for such studies is that the genetic isolate is representative for the general population or disease-specific study populations. This is indeed the case since we showed that in selected subjects with COPD from the general population, we can replicate the associations found in the young genetic isolate for a substantial part. Thus, we are able to translate findings in a genetic isolate to the general population, but correct and comparable phenotyping of the study populations is still crucial to replicate association between populations. We were unable to replicate results of any of the SNPs

in subjects with GOLD stage ≥ I from the Vlagtwedde/Vlaardingen population. When looking more closely, it appeared that the GRIP COPD patients with GOLD stage ≥ I had more severe COPD, i.e. lower lung function and more symptoms, than COPD patients with similar stage of disease in the Vlagtwedde/Vlaardingen population (table 1). The most important explanation for this difference is that we included COPD patients based on VC values in the Vlagtwedde/Vlaardingen population, whereas they were selected on the basis of FVC values in GRIP. Due to the maneuver of measurement, a VC is generally larger than an FVC and therefore the inclusion based on the cut-off value of $FEV_1/FVC< 70\%$ leads to inclusion of more patients in Vlagtwedde/Vlaardingen. A more strict definition of COPD (GOLD stage ≥ II) in Vlagtwedde/Vlaardingen gives a phenotypically better comparison between patients in GRIP and Vlagtwedde/Vlaardingen (table 1). Indeed, when analyzing subjects with subjects with GOLD ≥ II from Vlagtwedde/Vlaardingen population, SNPs *ADAM33* ST+5, *TGFβ1* C-509T and *TGFB1* Leu10Pro were significantly associated with FEV_1/VC, as they were in the GRIP GOLD ≥ I COPD patients.

Several explanations may exist for the lack of replication for *SFTPA1*, *SFTPA2* and *SFTPD* SNPs with FEV_1/VC in the Vlagtwedde/Vlaardingen COPD GOLD stage ≥ II. First, the original GRIP findings on these genes could be false-positive. Indeed, multiple (though correlated) outcomes and SNPs were studied in GRIP. Another, more biological explanation, for the lack of replication may be that the prevalence of certain alleles in genetically isolated populations differs from a general population as a result of genetic drift and founder effects. However, genotype frequencies were not statistically different between the two populations (see table 4). Yet a third explanation may be that differences exist in characteristics between the study populations. The GRIP population had more severe COPD and was slightly older than the Vlagtwedde/ Vlaardingen COPD population (see table 1). In addition, differences in environment may affect the lack of replication of the *SFTP* genes. The genetically isolated population shares the same environment, similar socio-economic status, and the same general practitioners. We cannot rule out that the COPD patients in GRIP have a higher prevalence of chronic bronchitis and airway disease whereas the airway obstruction in the Vlagtwedde/Vlaardingen population may have been caused by emphysema [28-30]. Further research is needed to separately assess these phenomena, since CT scans are necessary, which we unfortunately do not have of these patients.

In conclusion, this study provides two important messages. Firstly, we found significant effects of SNPs on the severity of COPD, i.e. level of lung function in patients with established COPD in a relatively small genetically isolated population with a large pedigree structure. Secondly, we replicated many of these associations in COPD patients selected from the general population on the condition that they were phenotypically similar. These findings are important since more severe airway obstruction is associated with progression and mortality of COPD. Future studies using this genetic isolate should focus

on progression of COPD, since this population seems to be highly usable to determine genetic risk factors for severity of airway obstruction in established COPD that can be translated to the general population.

Table 4: Genotype frequencies of the significant SNPs in the GRIP population GOLD stage ≥ I compared to the genotype frequencies in the Vlagtwedde/Vlaardingen population GOLD stage ≥ I and comparison of genotype frequencies between GOLD stage ≥ II populations

SNP and genotype		GRIP GOLD ≥ I N=106 N (%)	Vla/Vla GOLD ≥ I N=351 N (%)	P value	GRIP GOLD ≥ II N=70 N (%)	Vla/Vla GOLD ≥ II N=167 N (%)	P value
ADAM33	AA	15 (14.4)	63 (18.1)	0.682	12 (17.1)	27 (16.5)	0.475
ST+5	AG	53 (51.0)	171 (49.1)		31 (44.3)	86 (52.4)	
	GG	36 (34.6)	114 (32.8)		27 (38.6)	51 (31.1)	
TGFβ1	GG	46 (44.7)	187 (54.0)	0.100	30 (42.9)	94 (58.0)	0.049
C-509T	GA	45 (43.7)	137 (39.6)		32 (45.7)	60 (37.0)	
	AA	12 (11.6)	22 (6.4)		8 (11.4)	8 (4.9)	
TGFβ1	AA	35 (35.4)	131 (40.4)	0.655	23 (34.8)	68 (45.3)	0.341
Leu10Pro	AG	50 (50.5)	149 (46.0)		35 (53.0)	65 (43.3)	
	GG	14 (14.1)	44 (13.6)		8 (12.1)	17 (11.3)	
SFTPA1	GG	70 (72.9)	258 (78.4)	0.513	47 (73.4)	123 (79.9)	0.103
Leu50Val	GC	20 (20.8)	53 (16.1)		15 (23.4)	20 (13.0)	
	CC	6 (6.3)	18 (5.5)		2 (3.1)	11 (7.1)	
SFTPA2	GG	70 (66.6)	250 (72.7)	0.480	45 (64.3)	117 (72.7)	0.328
Pro91Ala	GC	32 (30.5)	87 (25.3)		22 (31.4)	41 (25.5)	
	CC	3 (2.9)	7 (2.0)		3 (4.3)	3 (1.9)	
SFTPD	TT	40 (39.2)	105 (30.9)	0.229	24 (35.8)	44 (27.5)	0.419
Met11Thr	TC	42 (41.2)	170 (50.0)		30 (44.8)	85 53.1)	
	CC	20 (19.6)	65 (19.1)		13 (19.4)	31 (19.4)	
SFTPD	AA	37 (37.0)	111 (32.9)	0.730	27 (40.3)	54 (34.6)	0.664
Ala160Thr	AG	47 (47.0)	165 (49.0)		30 944.8)	73 (46.8)	
	GG	16 (16.0)	61 (18.1)		10 (14.9)	29 (18.6)	

Abbreviations: GRIP Genetic Research in Isolated populations; Vla/Vla Vlagtwedde/Vlaardingen; GOLD Global initiative for Obstructive Lung Diseases;ADAM33 A Disintegrin and Metalloprotease 33; TGFβ1 Transforming Growth Factor β1; SFTP Surfactant Protein

ACKNOWLEDGEMENTS

We thank L. Testers for his help in logistics of the fieldwork and DNA collection, and M. Farenhorst, Th. van Hoogdalem J. Post, A. Verbokkem, and K. Vink-Klooster for collecting the lung function data. We would like to thank P. Veraart, H. Kornman and E. Boeren for their contribution to genealogical research. The Dutch Asthma Foundation funded collection of lung function data. C.C. van Diemen is assigned through the Dutch Asthma Foundation (NAF3.2.02.51). GRIP is supported by grants from the Netherlands Organization for Scientific Research (NWO: Pionier grand Dr C.M. van Duijn), the Center for Medical Systems Biology (CMSB). We would like to thank all participants of the GRIP study for their cooperation as well as the neurologists and general practitioners that made this work possible. All research assistants of the GRIP study are acknowledged for their help in data collection.

REFERENCES

1. WHO. World Health Report 2002. Available from: URL: http://www.who.int/whr/2002
2. Murtagh E, Heaney L, Gingles J, *et al.* The prevalence of obstructive lung disease in a general population sample: the NICECOPD study. *Eur J Epidemiol* 2005;20(5):443-53.
3. Ekberg-Aronsson M, Pehrsson K, Nilsson JA, *et al.* Mortality in GOLD stages of COPD and its dependence on symptoms of chronic bronchitis. *Respir Res* 2005;6:98.
4. Sin DD, Man SF. Chronic obstructive pulmonary disease as a risk factor for cardiovascular morbidity and mortality. Proc Am Thorac Soc 2005;2(1):8-11.
5. Hospers JJ, Postma DS, Rijcken B, *et al.* Histamine airway hyper-responsiveness and mortality from chronic obstructive pulmonary disease: a cohort study. *Lancet* 2000;356(9238):1313-7.
6. Pardo LM, Mackay I, Oostra B, *et al.* The effect of genetic drift in a young genetically isolated population. *Ann Hum Genet* 2005;69(Pt 3):288-95.
7. Aulchenko YS, Heutink P, Mackay I, *et al*. Linkage disequilibrium in young genetically isolated Dutch population. *Eur J Hum Genet* 2004;12(7):527-34.
8. Njajou OT, Vaessen N, Joosse M, *et al.* A mutation in SLC11A3 is associated with autosomal dominant hemochromatosis. *Nat Genet* 2001;28(3):213-4.
9. Liu F, Elefante S, van Duijn CM, *et al*. Ignoring distant genealogic loops leads to false-positives in homozygosity mapping. *Ann Hum Genet* 2006;70(Pt 6):965-70.
10. Liu F, Rias-Vasquez A, Sleegers K, *et al.* A genomewide screen for late-onset Alzheimer disease in a genetically isolated Dutch population. *Am J Hum Genet* 2007;81(1):17-31.
11. Boezen HM, Vonk JM, van Aalderen WM, *et al.* Perinatal predictors of respiratory symptoms and lung function at a young adult age. *Eur Respir J* 2002;20(2):383-90.
12. Quanjer PH, Tammeling GJ, Cotes JE, *et al.* Lung volumes and forced ventilatory flows. Report Working Party Standardization of Lung Function Tests, European Community for Steel and Coal. Official Statement of the European Respiratory Society. *Eur Respir J Suppl* 1993;16:5-40.
13. van Diemen CC, Postma DS, Vonk JM, *et al.* A disintegrin and metalloprotease 33 polymorphisms and lung function decline in the general population. *Am J Respir Crit Care Med* 2005;172(3):329-33.
14. van Diemen CC, Postma DS, Vonk JM, *et al.* Decorin and TGF-beta1 polymorphisms and development of COPD in a general population. *Respir Res* 2006;7:89.
15. Boerwinkle E, Chakraborty R, Sing CF. The use of measured genotype information in the analysis of quantitative phenotypes in man. I. Models and analytical methods. *Ann Hum Genet* 1986;50(Pt 2):181-94.
16. Aulchenko YS, de Koning DJ, Haley C. Genomewide rapid association using mixed model and regression: a fast and simple method for genomewide pedigree-based quantitative trait Loci association analysis. *Genetics* 2007;177(1):577-85.
17. Gilmour AB, Gogel BJ, . ASReml user guide release 1.0. NSW Agriculture, Orange, Australia ed. 2002.
18. From the Global Strategy for the Diagnosis, Management and Prevention of COPD, Global Initiative for Chronic Obstructive Lung Disease (GOLD) 2006. Available from: http://www.goldcopd.org.

19. Gosman MM, Boezen HM, van Diemen CC, *et al.* A disintegrin and metalloprotease 33 and chronic obstructive pulmonary disease pathophysiology. *Thorax* 2007;62(3):242-7.

20. van Diemen CC, Koeter GH, Timens W, *et al.* ADAM33 SNPs are associated with excess FEV1 decline in severe early onset COPD. *Eur Respir J* 2006;28[Suppl 50], 582s-583s.

21. Celedon JC, Lange C, Raby BA, *et al.* The transforming growth factor-beta1 (TGFB1) gene is associated with chronic obstructive pulmonary disease (COPD). *Hum Mol Genet* 2004;13(15):1649-56.

22. Su ZG, Wen FQ, Feng YL, *et al*. Transforming growth factor-beta1 gene polymorphisms associated with chronic obstructive pulmonary disease in Chinese population. *Acta Pharmacol Sin* 2005;26(6):714-20.

23. Wu L, Chau J, Young RP, *et al.* Transforming growth factor-beta1 genotype and susceptibility to chronic obstructive pulmonary disease. *Thorax* 2004;59(2):126-9.

24. Ogawa E, Ruan J, Connett JE, *et al*. Transforming growth factor-beta1 polymorphisms, airway responsiveness and lung function decline in smokers. *Respir Med* 2007;101(5):938-43.

25. Yoon HI, Silverman EK, Lee HW, *et al*. Lack of association between COPD and transforming growth factor-beta1 (TGFB1) genetic polymorphisms in Koreans. *Int J Tuberc Lung Dis* 2006;10(5):504-9.

26. Guo X, Lin HM, Lin Z, *et al*. Surfactant protein gene A, B, and D marker alleles in chronic obstructive pulmonary disease of a Mexican population. *Eur Respir J* 2001;18(3):482-90.

27. van Diemen CC, Postma DS, Vonk JM, *et al*. Polymorphisms in surfactant proteins and FEV1 decline and development of COPD in the general population. *Eur Respir J* 2006;28[Suppl 50], 143s-144s.

28. DeMeo DL, Hersh CP, Hoffman EA, *et al.* K. Genetic Determinants of Emphysema Distribution in the National Emphysema Treatment Trial. *Am J Respir Crit Care Med* 2007; 176(1):42-8.

29. Martinez FJ, Foster G, Curtis JL, *et al*. Predictors of mortality in patients with emphysema and severe airflow obstruction. *Am J Respir Crit Care Med* 2006 Jun 15;173(12):1326-34.

30. Martinez FJ, Curtis JL, Sciurba F, *et al*. Gender Differences in Severe Pulmonary Emphysema. *Am J Respir Crit Care Med* 2007;176(3):243-52.

Online supplement for paper:

Novel strategy to identify genetic risk factors for COPD severity: genetically isolated populations

Cleo C van Diemen, Dirkje S Postma, Yurii S Aulchenko, Pieter JLM Snijders, Ben A Oostra, Cornelia M van Duijn, H Marike Boezen

Table E1: Characteristics of the GRIP study populations by GOLD staging
Data are presented as median (range).

	GRIP population according to COPD status		
	≥GOLD I n=106	≥GOLD II n=70	No COPD n=51
Age	65.5 (37-84)	66.0 (43-82)	61.0 (28-83)
Gender m/f	53/53	37/33	18/33
Smoking status			
Never smoker, %	3.8	2.9	9.8
Ex-smoker, %	34.0	37.1	33.3
Current smoker, %	58.5	57.1	47.1
Missing smoking history, %	3.8	2.9	9.8
Pack-years smoking	32.3 (0-120)	39 (0-120)	28 (0-74)
FEV_1 % predicted	71.2 (26.4-132.8)	63.9 (26.4-79.0)	95.2 (64.1-129.3)
FEV_1/FVC, %	57.7 (27.7-69.2)	53.1 (27.7-68.9)	76.0 (70.1-90.1)

CHAPTER 12

Summary, General Discussion & Future Perspectives

Summary

Introduction

Chronic Obstructive Pulmonary Disease (COPD) is one of the leading causes of morbidity and mortality worldwide. Central to the definition of COPD is a reduced level of lung function and this provides the most important and robust phenotype of COPD. Smoking is the major risk factor for COPD development. However, not all smokers develop COPD, indicating a genetic predisposition for this disease. This thesis deals with genetic factors influencing COPD development and accelerated decline in lung function in the general population, and the severity of the disease in patients with established COPD.

Genetic epidemiology of lung function

In ***chapter 2***, several study designs have been discussed that are being used in genetic epidemiology to study genetic determinants of reduced lung function. *Family and twin studies* provided the first evidence for familial aggregation of lung function measures, indicating that genetic factors are important for level of lung function. Difficulties in obtaining large sample sizes are a problem in this type of study design and especially in genetic epidemiology of COPD such studies are difficult to perform due to the late onset of disease. *Genome wide linkage analysis* is an approach to identify which chromosomal areas are involved in disease. This method compares the inheritance of disease with segregation of genetic markers in families with multiple affected members. If the disease is co-inherited with the marker, this suggests that a disease susceptibility gene is located close to the marker on that chromosome. Linkage analyses for lung function have provided multiple chromosomal regions thus far, but only a few candidate genes have been identified. In *case-control association studies*, a large number of candidate genes, chosen for their biological role in the pathogenesis in COPD, have been associated with lung function or COPD. However, for every studied gene there is (at least) one study in which the results contradict results of other studies. Replication of results in different studies is difficult due to population stratification, ethnic differences, and small numbers of cases and controls that are usually included in studies. *Longitudinal cohort studies* that evaluate genetic influences on lung function decline in combination with environmental factors are important for identifying subjects at risk for COPD. However, since such studies are scarce, limited knowledge on the effect of genes on the course of lung function on a general population level is available. Studies in *isolated populations* can be used in genetic epidemiology to identify candidate regions in a relatively small population because of their decreased genetic heterogeneity, increased linkage disequilibrium and more

homogeneous environmental exposures. Studying isolated populations has the advantage that data collection for phenotyping (e.g. lung function measurements) and genotyping can usually be performed in a logistically small area, with additionally low cost.
Overall, every type of design has its restrictions, varying e.g. from high cost to relatively low costs, and hypothesis-generating to hypothesis-testing. Choices about the most appropriate study design, or the population under study are usually made dependent of the goal of the study, e.g. whether one sets out to confirm a previously found association in a different population, or one searches for new genes or gene-pathways that have not been studied so far.

A Disintegrin And Metalloprotease 33

A Disintegrin and Metalloprotease 33 (ADAM33) has been identified by positional cloning as a susceptibility gene for asthma, and especially in asthma with bronchial hyperresponsiveness (BHR). Single nucleotide polymorphisms (SNPs) in this gene have been associated with excess decline in the Forced Expiratory Volume in 1 second (FEV_1) in asthmatics and with BHR [1]. We hypothesized that *ADAM33* may not only be associated with excess FEV_1 decline and BHR in asthmatics, but that it could constitute a broader phenomenon that might be observed in the general population and COPD as well. Therefore, we performed 3 studies to assess the effect of genetic variation in *ADAM33* on lung function, COPD and COPD severity.

In ***chapter 3***, we describe the role of SNPs in *ADAM33* on lung function decline in 1,390 subjects of the Vlagtwedde/Vlaardingen cohort who have been followed up from 1965 to1990. We analyzed 8 SNPs in *ADAM33* and used linear mixed effects models to analyze FEV_1 decline according to genotype, adjusting for sex, age, first available FEV_1 and pack-years of smoking. The mean unadjusted decline in the cohort was 18.6 ml/per year. Individuals homozygous for minor alleles of SNPs S_2 and Q-1 and heterozygous for SNP S_1 had a significantly excess decline in FEV_1 of respectively 4.9, 9.6 and 3.6 ml/y compared to the wild type allele. In addition, we found a significantly higher prevalence of the minor alleles of SNPs F+1, S_1, S_2 and T_2 in subjects with COPD (defined according to the Global Initiative for Obstructive Lung Diseases (GOLD) stage ≥ II at the last 2 surveys) compared to subjects without COPD (respectively 61.3 vs. 53.3%; 22.9 vs. 14.3%; 51.9 vs. 41.7%; 30.6 vs. 23.6%).

We conclude that SNPs in *ADAM33* are associated with accelerated lung function decline in the general population and they also confer a risk to develop COPD.
Since *ADAM33* has previously been associated with BHR in asthma [1-5] and the vast majority of COPD patients displays BHR [6], *ADAM33* may also play a role in the severity of BHR within COPD patients. In ***chapter 4***, we therefore studied the effect of *ADAM33* SNPs

on BHR severity and inflammation in sputum and bronchial biopsies in 114 patients with mild to moderate COPD (GOLD stages II and III) from the GLUCOLD study. COPD patients with the *ADAM33* ST+5 AA genotype had more severe BHR, higher numbers of sputum inflammatory cells and higher numbers of $CD8^+$ cells in bronchial biopsies than patients with the GG-genotype. $CD8^+$ cell numbers in bronchial biopsies were significantly higher in subjects carrying the minor allele of SNP S_2 or T_2 than those with the wild-type.
This study shows that *ADAM33* plays a role not only in the development of COPD, but also affects severity of BHR and inflammation in patients with established COPD.
As *ADAM33* appeared to be associated with phenotypes reflecting COPD severity such as BHR -by itself a risk factor for accelerated lung function decline in COPD patients- *ADAM33* SNPs may also play a role in excess FEV_1 decline within COPD patients. Hence, we investigated *ADAM33* SNPs on FEV_1 decline in 79 patients with severe COPD with and without α1-antitrypsin (AAT)-deficiency with longitudinal lung function data prior to lung transplantation (LTX) (***chapter 5***). AAT-deficiency is a genetic constitution, thought to be the main factor for excess lung function decline and development of COPD at a young age in carriers of AAT-deficiency, even without cigarette smoking. We found that the *ADAM33* ST+5 and T_2 SNPs were associated with excess FEV_1 decline, adjusted for sex, AAT-deficiency, pack-years smoking and baseline FEV_1. Stratified analyses according to AAT-deficiency showed comparable results.

This study shows that *ADAM33* SNPs are predictive of excess FEV_1 decline in severe COPD, and interestingly also in patients with accelerated FEV_1 loss due to AAT-deficiency.
It is clear from genetic studies that *ADAM33* plays a role in the origin and progression of lung diseases. The focus of current studies should lie in gaining knowledge on the function of ADAM33 and its role in respiratory disorders. ADAM33 protein and mRNA expression has previously been demonstrated in fibroblasts, airway smooth muscle and mesenchymal cells from asthmatics and recently to a smaller extent in epithelial cells from both asthmatics and controls [7-11]. However, so far it has remained unclear if and where *ADAM33* is expressed in COPD lung tissue. In ***chapter 5***, we have therefore performed immunohistochemical staining of ADAM33 on lung tissue of the 79 COPD patients from the LTX study, as well as control tissue. We found expression in endothelium, airway smooth muscle, alveoli, inflammatory cells and bronchial epithelium, predominantly in the basal part of the epithelium, with antibodies directed against the cytoplasmic and catalytic domain of ADAM33. There were no differences in level or location of ADAM33 expression between tissue samples from control, AAT-deficient and non-AAT-deficient subjects. The expression pattern and intensity was similar using the different antibodies. These results show that it is not likely a difference in level or location of ADAM33 expression that confers a risk for lung disease, but merely a functional change in ADAM33.

MATRIX METALLOPROTEASES, THEIR TISSUE INHIBITORS AND GENES INVOLVED IN TISSUE REPAIR

A generally accepted underlying mechanism in COPD pathology is a disturbed balance between matrix metalloproteinases (MMPs) and Tissue Inhibitors of MMPs (TIMPs) in lung tissue of COPD patients, causing degradation of extracellular matrix (ECM). In addition, repair of ECM is impaired in COPD tissue, which contributes to accelerated lung function decline. These processes of disturbed protease-antiprotease balances and repair may be genetically determined due to SNPs in the genes involved in these processes.
In ***chapter 6***, we assessed the association of SNPs in *MMP* and *TIMP* genes with COPD and accelerated FEV_1 decline in the Vlagtwedde/Vlaardingen cohort. We analyzed SNPs in *MMP1* (-1607 G/GG), *MMP2* (-1306 C/T), *MMP9* (3 haplotype tagging SNPs), *MMP12* (-82 A/G and Asn357Ser) and the X-linked gene*TIMP1* (Phe124Phe and Ile158Ile). All significant associations were repeated in an independent general population cohort, the Doetinchem cohort (n=1,152). In the Vlagtwedde/Vlaardingen cohort, *MMP2* -1306 TT genotype carriers had excess FEV_1 decline (-4.0 ml/yr, p=0.03) compared to wild type carriers. *TIMP1* Ile158Ile predicted significant excess FEV_1 decline in males and females. *TIMP1* Phe124Phe predicted significant excess FEV_1 decline in males only, which was replicated in the Doetinchem cohort. The *MMP2* and *TIMP1* Ile158Ile associations were not replicated. This study shows that *TIMP1* Phe124Phe contributes to excess FEV_1 decline in two independent prospective cohorts. Further research is needed to assess the functionality of *TIMP1* Phe124Phe Genetic variations in *MMPs,* although they have functional effects on MMP expression levels, evidently do not contribute to FEV_1 decline in the general population.

Decorin, an ECM proteoglycan, and Transforming Growth Factor-β_1 (TGF-β_1) are both involved in lung ECM turnover. Decorin and TGF-β_1 expression are decreased respectively increased in COPD lung tissue. Interestingly, they act as each other's feedback regulator, as depicted in figure 1 below.

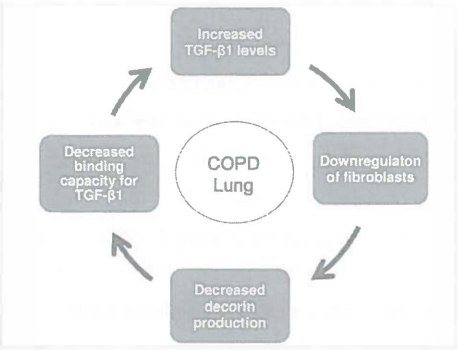


Figure 1: Loop of a sustained surplus of TGF-β1 and increasingly reduced decorin level in COPD lung tissue

We hypothesized that the reciprocal regulation of the *TGF-β$_1$* and *decorin* genes is disturbed in COPD due to a genetic mutation in one or both of these genes. In ***chapter 7***, we have tested this hypothesis by investigating three SNPs in *TGF-β$_1$* and five SNPs in *decorin* on the development of COPD and on lung function decline in the Vlagtwedde/Vlaardingen cohort. We tested five SNPs in *decorin* (3'untranslated region (UTR) and four intron SNPs) and three SNPs in *TGF-β$_1$* (3'UTR (rs6957), C-509T (rs1800469) and Leu10Pro (rs1982073)), and their haplotypes. We found a significantly higher prevalence of carriers of the minor allele of the *TGF-β1* rs6957 SNP (p=0.001) in subjects with COPD (GOLD stage ≥ II). Additionally, we found a significantly lower prevalence of the haplotype with the major allele of rs6957 and minor alleles for rs1800469 and rs1982073 SNPs in *TGF-β$_1$* in subjects with COPD (p=0.030), indicating that this association is due to the rs6957 SNP. *TGF-β$_1$* SNPs were not associated with FEV_1 decline. SNPs in *decorin,* and haplotypes constructed of both *TGF-β$_1$* and *decorin* SNPs were not associated with development of COPD or with FEV_1 decline.

Contrary to our hypothesis, we were not able to identify the *decorin* gene as a genetic risk factor for the development of COPD. Consequently, SNPs in *decorin* do not seem to underlie a disturbed regulation of this gene and *TGF-β$_1$* resulting in COPD, nor can they be held responsible for the development of COPD and decline in FEV_1 in the general population. We found that *TGF-β$_1$* SNPs are associated with the development of COPD but not with accelerated lung function decline or other lung function measures in the general population. Together with previous findings, this study establishes the *TGF-β$_1$* gene as a risk factor for the development of COPD.

GENES INVOLVED IN OXIDATIVE STRESS

COPD may develop due to excess oxidants in the lungs that are not properly counterbalanced by either exogenous or endogenous antioxidants or antioxidant enzymes. This leads to oxidative stress which is associated with inflammation of the airways.

The glutathione system is the major antioxidant system in the lungs that protects the lungs against inhaled oxidants. The Glutathion S-Transferase (GST) enzymes are key enzymes in this oxidant-antioxidant balance and GSTs can detoxify certain components of cigarette smoke. In ***chapter 8***, we analyzed the effect of deletion alleles of *GSTM1* and *GSTT1* and 2 SNPs (Ile105Val and Ala114Val) in *GSTP1* on the course of lung function and COPD development in the Vlagtwedde/Vlaardingen population. We found that subjects heterozygous for the *GSTT1* null allele had a significantly higher risk to develop COPD (odds ratio 2.03, p=0.002) and also had significant excess FEV_1 decline (-17.3 vs. -20.7 ml/yr) compared to subjects with the homozygous wild type null allele. Subjects

homozygous for the *GSTM1* null allele had excess FEV_1 decline compared to carriers of at least one wild type allele (-16.5 vs. -18.9 ml/yr). Stratified analyses according to smoking (never vs. ever) and gender showed a significant association of the *GSTM1* null allele with excess FEV_1 decline in male ever smokers, but not in male never smokers. *GSTT1* was significantly associated with FEV_1 decline in males only, irrespective of smoking. The *GSTM1* and *GSTT1* null alleles were not associated with excess FEV_1 decline in females, also not when stratified according to smoking.
In conclusion, the *GSTM1* and *GSTT1* null alleles confer risk for excess FEV_1 decline especially in male smokers. The *GSTT1* null allele additionally confers a risk to develop COPD. The *GSTP1* SNPs do not appear to be important in lung function decline or COPD development in the general population.

Oxidative stress can be counterbalanced by different mechanisms. Superoxide Dismutases (SODs) act in a different pathway than the glutathion system and prevent the elevation of oxidative stress by degradation of superoxide anions. In ***chapter 9***, we investigated the association of SNPs located in *SOD-2* and *SOD-3* with development of bronchial hyperresponsiveness (BHR) and COPD, and the longitudinal course of FEV_1 over time in the Vlagtwedde/Vlaardingen cohort. We genotyped 2 haplotype-tagging SNPs (Ala16Val (rs4880) and C7693T (rs2842958)) in *SOD-2* and 3 nonsynonymous SNPs (Ala40Thr (rs2536512), Arg213Gly (rs1799895) and Phe131Cys) in *SOD-3*. The C7693T SNP, located in intron 3 of *SOD-2,* was significantly associated with the presence of COPD and BHR (defined as $PC_{10} \leq 8$ mg/ml of histamine at the last survey) in the total population. The T/T genotype for this SNP and the Val/Val genotype for the *SOD-2* Ala16Val substitution were risk factors for BHR in individuals without COPD. The *SOD-3* Arg213Gly substitution was associated with slower FEV_1 decline in never-smokers, e.g. 9.3 ml less decline per year in carriers of the Arg/Gly genotype while compared to the wild type.

We concluded from this study that the two *SOD-2* SNPs are risk factors for BHR in absence of COPD while the *SOD-2* C7693T SNP is a risk factor for COPD in the total population. Finally, the *SOD-3* Arg213Gly SNP is protective for FEV_1 decline in never-smokers.

GENETIC AND LIPIDOMIC ASSOCIATIONS OF SURFACTANT COMPONENTS WITH LUNG FUNCTION

Pulmonary surfactant functions in the first response to microorganisms in the lung, regulation of inflammation and structure of alveoli. It consists of lipids and specific surfactant proteins. A dysfunctional surfactant may be involved in COPD development, lower lung function, faster lung function decline and related traits. ***Chapter 10*** describes our study to determine risk factors for COPD in both the protein and lipid part of the pulmonary surfactant. We analyzed the effect of SNPs in the *surfactant protein* (*SFTP*)

genes *SFTPA1*, *SFTPA2*, *SFTPB* and *SFTPD* on lung function decline, development of COPD and on well-known risk factors of COPD, i.e. BHR and blood eosinophilia in the Vlagtwedde/Vlaardingen cohort. SNPs *SFTPA1* Leu50Val and *SFTPD* Met11Thr were significant predictors of excess FEV_1 decline and COPD development. *SFTPD* Met11Thr was associated with blood eosinophilia, a risk factor for accelerated FEV_1 decline.
In addition to the genetic study, we analyzed surfactant lipids and their degradation products in sputum of 17 ex-smoking COPD patients and 11 ex-smoking healthy controls and tested their association with lung function parameters. COPD patients had significantly lower phosphatidylcholine and higher lysophosphatidylcholine levels than controls. Accordingly, phosphatidylcholine levels were positively and lysophosphatidylcholine levels were negatively associated with lung function parameters. We thus show for the first time that genetic factors and lipid composition in surfactant can negatively affect lung function, thereby contributing to COPD development.

IDENTIFICATION OF GENETIC RISK FACTORS FOR COPD SEVERITY IN A GENETICALLY ISOLATED POPULATION

Genetic determinants for COPD are difficult to study, since COPD develops at later ages and involves multiple genetic and environmental factors. Studies using genetically isolated populations with limited genetic variation may be useful in COPD genetics, but are thus far lacking. In ***chapter 11***, we studied the association between SNPs in candidate genes for COPD and severity of COPD in a genetically isolated population. We collected spirometry and questionnaire information from 157 subjects with a doctor's diagnosed COPD from the larger Genetic Research in Isolated Population (GRIP) study. We genotyped SNPs in 13 candidate genes: *ADAM33*, *TGF-β1*, *MMP1*, *MMP2*, *MMP9*, *MMP12*, *TIMP1*, *SFTPA1*, *SFTPA2*, *SFTPB*, *SFTPD* and *GSTP1* and studied their associations with FEV_1, FVC and FEV_1/FVC levels within COPD patients. Using spirometry, we confirmed COPD GOLD stage ≥ I in 106 subjects, who were members of an extended pedigree including 6,175 people. SNPs in *ADAM33*, *TGF-β1*, *SFTPA1*, *SFTPA2* and *SFTPD* were significantly associated with FEV_1/Forced Vital Capacity (FVC) level. We replicated the significant associations from the GRIP population in COPD patients selected from the general population of Vlagtwedde/Vlaardingen. We found that the same *ADAM33* and *TGF-β1* SNPs were associated with FEV_1/FVC in COPD patients from the general population. In this study, we have shown for the first time that a genetic isolate can be used to determine genetic risk factors for lung function level, which can be replicated in COPD patients from an independent population.

MAIN CONCLUSIONS

The chapters in this thesis describe multiple studies in which many genes and SNPs have been tested in different study populations with diverse phenotypes. To make the results of this thesis more comprehensible, table 1 presents an overview of all genes and SNP studied, the (lack of) associations found, the cohorts in which these associations were found and with which phenotype.

The main conclusions of this thesis are:

1. *ADAM33* SNPs are associated with excess decline in lung function and development of COPD in the general population and with severity of COPD in patients with established COPD.
2. The X-chromosomal *TIMP1* gene is a risk factor for excess FEV_1 decline in males in two independent cohorts of the general population.
3. *TGF-β1* SNPs play a role in development of COPD in the general population, and with the severity of airway obstruction in COPD patients.
4. Null alleles of the antioxidant enzymes *GSTM1* and *GSTT1* interact with smoking and gender, and affect FEV_1 decline in the general population.
5. SNPs in the antioxidant enzymes *SOD-2* and *SOD-3* are associated with BHR and FEV_1 decline in the general population.
6. Surfactant protein SNPs are associated with FEV_1 decline, COPD development, and blood eosinophilia in the general population, and with severity of airway obstruction in COPD patients.
7. Surfactant lipid degradation in sputum is associated with lower lung function.
8. The SNPs studied in *decorin*, *MMP1*, *MMP9*, *MMP12*, *SFTPB*, *GSTP1* do not appear to be important in excess FEV_1 decline and COPD development on a general population level.
9. Genetically isolated populations can be used to identify genetic risk factors for COPD severity in a relatively small sample size that can be replicated in the general population.

Table 1: Genes, cohorts and SNPs studied in this thesis and the associations found

Gene	SNPs genotyped rs numbers	alternative names	Study population	SNPs and phenotypes
ADAM33	rs17548913	*ADAM33*F+1	*Vla/Vla*	excess FEV_1 decline (Q-1, S1, S2); COPD development (F+1, S2, T1, T2); lower levels of FEV_1/VC in COPD patients GOLD ≥ stage II (ST+5)
	rs17548907	*ADAM33*Q-1		
	rs3918396	*ADAM33*S1		
	rs528557	*ADAM33*S2		
	rs597980	*ADAM33*ST+5	*GLUCOLD*	severity of inflammation in sputum and bronchial biopsies (ST+5, S2, T2); severity of BHR (ST+5)
	rs2280091	*ADAM33*T1		
	rs2280090	*ADAM33*T2		
	rs2787094	*ADAM33*V4	*LTX*	excess FEV_1 decline (ST+5)
			GRIP	lower levels of FEV_1/FVC in COPD patients GOLD ≥ stage I (ST+5)
MMP1	rs1799750	*MMP1*G-1607GG	*Vla/Vla*	no association
			Doetinchem	no association
			GRIP	no association
MMP2	rs243865	*MMP2*C-1306T	*Vla/Vla*	excess FEV_1 decline
			Doetinchem	no association
			GRIP	no association
MMP9	rs3918278	*MMP9*_rs3918278	*Vla/Vla*	no association
	rs6065912	*MMP9*_rs6065912	*GRIP*	no association
	rs8113877	*MMP9*_rs8113877		
MMP12	rs2276109	*MMP12*A-82G	*Vla/Vla*	no association
	rs652438	*MMP12*Asn357Ser	*GRIP*	no association
TIMP1	rs11551797	*TIMP1*Phe124Phe	*Vla/Vla*	excess FEV_1 decline in males and females (*TIMP1*Phe124Phe, *TIMP1*Ile158Ile)
	rs4898	*TIMP1*Ile158Ile		
			Doetinchem	excess FEV_1 decline in males (*TIMP1*Phe124Phe)
			GRIP	no association
TGF-β$_1$	rs1800469	*TGFβ1*C-509T	*Vla/Vla*	COPD development (*TGFβ1*3UTR); lower levels of FEV_1/VC in COPD patients GOLD ≥ stage II (*TGFβ1*C-509T, *TGFβ1*Leu10Pro)
	rs1982073	*TGFβ1*Leu10Pro		
	rs6957	*TGFβ1*3UTR		
			GRIP	lower levels of FEV_1/FVC in COPD patients GOLD ≥ stage I (*TGFβ1*C-509T, *TGFβ1*Leu10Pro)
Decorin	rs11106030	-	*Vla/Vla*	no association
	rs741212	-		
	rs516115	-		
	rs3138241	-		
	rs1803343	-		

Gene	SNPs genotyped rs numbers - alternative names		Study population	SNPs and phenotypes
GSTP1	rs1695 rs1138272	*GSTP1*Ile105Val *GSTP1*Ala114Val	*Vla/Vla*	no association
			GRIP	no association
GSTM1	-	Null allele	*Vla/Vla*	excess FEV_1 decline in male smokers
GSTT1	-	Null allele	*Vla/Vla*	excess FEV_1 decline; COPD development
SOD-2	rs4880 rs2842958	*SOD2*Ala16Val *SOD2*C7693T	*Vla/Vla*	BHR in absence of COPD (*SOD2*Ala16Val, *SOD2*C7693T); COPD (*SOD2*C7693T)
SOD-3	rs2536512 rs1799895 -	*SOD3*Ala40Thr *SOD3*Arg213Gly *SOD3*Phe131Cys	*Vla/Vla*	excess FEV_1 decline (*SOD3*Phe131Cys)
SFTPA1	rs1059047 rs1136450 rs4253527	*SFTPA1*Val19Ala *SFTPA1*Leu50Val *SFTPA1*Arg219Trp	*Vla/Vla*	excess FEV_1 decline; COPD development; blood eosinophilia (*SFTPA1*Leu50Val)
			GRIP	lower levels of FEV_1/FVC in COPD patients GOLD ≥ stage I (*SFTPA1*Leu50Val)
SFTPA2	rs1059046 rs17886395 rs1965707	*SFTPA2*Asn9Thr *SFTPA2*Pro91Ala *SFTPA2*Ser140Ser	*Vla/Vla*	COPD development (*SFTPA2*Pro91Ala)
			GRIP	lower levels of FEV_1/FVC in COPD patients GOLD ≥ stage I (*SFTPA2*Pro91Ala)
SFTPB	rs1130866	*SFTPB*Ile131Thr	*Vla/Vla*	no association
			GRIP	no association
SFTPD	rs721917 rs2243639	*SFTPD*Met11Thr *SFTPD*Thr160Ala	*Vla/Vla*	excess FEV_1 decline (*SFTPD*Met11Thr); COPD development (*SFTPD*Met11Thr); blood eosinophilia (*SFTPD*Met11Thr, *SFTPD*Ala160Thr)
			GRIP	lower levels of FEV_1/FVC in COPD patients GOLD ≥ stage I (*SFTPD*Met11Thr, *SFTPD*Thr160Ala)

Abbreviations: SNP Single Nucleotide Polymorphism; Vla/Vla Vlagtwedde/Vlaardingen; GLUCOLD Groningen Leiden Universities Corticosteroids in Obstructive Lung Disease; LTX lung transplantation; GRIP Genetic Research in Isolated Population; COPD Chronic Obstructive Pulmonary Disease; GOLD Global initiative for Obstructive Lung Diseases; FEV1 Forced Expiratory Volume in 1 second; FVC Forced Expiratory Volume; BHR Bronchial Hyperresponsiveness

General discussion & future perspectives

INTRODUCTION

Many people suffer from COPD nowadays, and the disease is expected to increase even more worldwide [12]. The ultimate goal of performing genetic research is to find out why certain people develop the disease, identify those people and give them appropriate treatment at an early stage before the disease has progressed irreversibly. Since the major risk factor for COPD development is cigarette smoking, a complete ban on cigarette smoking would diminish the incidence of COPD a great deal. Still when cigarettes will be banned out completely, millions of people will suffer from cigarette smoke induced COPD for the upcoming years since the disease has a slowly progressive course and becomes only clinically apparent after many years of smoking and smoking cessation does not completely prevent this. Moreover, occupational exposures, environmental cigarette smoke and air pollution are responsible for approximately 10% of COPD cases, and with the increasing air pollution worldwide this number may grow in the future. Thus, the burden of COPD is high now and will remain so in the future decades. Therefore, identifying genetic risk factors will remain an important strategy to understand the mechanism of disease, and identify subjects who are at particular risk to develop this disease and/or are prone to (severe) progression.

Current standards define COPD by progressive loss of FEV_1, and thus longitudinal decline in FEV_1 will be the primary outcome variable for intervention studies aimed at preventing or reducing the loss of pulmonary function. It is therefore of importance to study genetic determinants for excess lung function decline as well as COPD development and COPD severity [13].

COPD is considered a complex genetic disease. The results from the studies presented in this thesis underline this and furthermore show that lung function level and decline in the general population is not influenced by one single gene, but by many genetic factors. We show that multiple genes are associated with COPD and with associated phenotypes, such as lung function level and decline, BHR and airway inflammation. We have used different cohorts to study COPD genetics from multiple points of view: from a general population based cohort consisting of respiratory healthy subjects to patients with very severe COPD. In addition, we have analyzed gene-environment interactions to assess which subjects are at particular risk for excess lung function decline or for development of COPD. The studies performed in the Vlagtwedde/Vlaardingen cohort and the different COPD patient populations thus provide different information and require different interpretations. Although it may seem to complicate matters to obtain associations with so many different phenotypes, the outcomes of these studies do supplement each other and provide a full

representation of how a gene may affect different aspects of COPD development and severity. Future studies should continue to take advantage of the possibilities these study populations provide. Therefore, in the following paragraphs, we will discuss some limitations of our studies and provide recommendations for future research in genetics of COPD.

USE OF DIFFERENT COHORTS AND INTERPRETATION OF RESULTS

General population

The majority of research described in this thesis has been carried out in the longitudinal general population based Vlagtwedde/Vlaardingen cohort, and thereby distinguishes itself from most other research in genetics of lung function. Case-control studies are generally used in genetic studies of COPD, and in these studies, COPD has already developed without acknowledging how the disease has evolved into clinically manifest disease. Other studies used selections from the general population with a specific phenotype or exposition, i.e. smokers with either a fast versus no decline in lung function, without phenotypic COPD [14-16]. Our studies in a longitudinal population-based cohort are particularly useful in that they allow determining which genetic risk factors are important in an early stage of lung function loss before clinical symptoms appear. In addition, they provide information that is applicable to the whole general population independent of disease. Moreover, our results may be used in an even broader perspective, since a low lung function is not only an important predictor for COPD development and progression, but also for cardiovascular diseases [17-20]. Recently, funding has been granted to study mortality within the participants of the Vlagtwedde/ Vlaardingen cohort. These mortality data are currently being collected, and it will be interesting to analyze genetic variation on the course of lung function, development of COPD and their relation to cause of mortality.

In the studies in Vlagtwedde/Vlaardingen population, we have merely looked at association of SNPs with excess decline in FEV_1. There are however other mechanisms that may lead to lung function impairment and COPD development, i.e. reduced maximal attained level of lung function, or abnormal early age of onset of (excess) lung function decline. Already *in utero*, altered branching of bronchi may take place, leading to impaired lung function in childhood. It has already been shown that some genes are associated with lung function in childhood in interaction with pre- or post natal smoke exposure, such as *GSTM1* and *ADAM33* [21-23]. Interesting future perspectives are to analyze whether SNPs affect more specifically these mechanisms in the general population. For such studies we need larger study populations with multiple lung function measurements from birth until the age where maximal attained lung function is reached (at approximately 35 years).

COPD patient studies

In the COPD patient populations, we have studied the genetic determinants of severity of lung function impairment rather than genetic determinants of initiation of lung function impairment. The genetics of COPD susceptibility and COPD severity may be different. Some genes may only contribute to disease susceptibility and other genes may contribute to disease severity, whereas again a third group of genes contribute to both. We have shown that some genes were indeed involved in the development of COPD or excess lung function decline in the general population, like *SFTPA, SFTPD and TIMP1*, whereas *ADAM33 and TGF-β1* were associated with features of lung function both in the general population and in COPD patient cohorts, thus in both disease development and disease progression. Another interesting finding presented in this thesis is that we have shown that SNPs in *ADAM33* may affect lung function decline even in the presence of AAT-deficiency. It is generally accepted that AAT-deficiency is the main underlying genetic cause of rapid FEV_1 decline in early onset COPD, but we show that other genetic factors may additionally play a role. Recently, another study has also shown that SNPs in *Interleukin-10* increase the risk for lower lung function, independent of AAT-deficiency [24]. It is important to determine whether polymorphisms are associated with COPD severity and progression, because it may help to target more aggressive and effective therapeutic approaches to groups with specific genetic background in the future.

ASSOCIATION WITH DIFFERENT PHENOTYPES

Lung function level and decline

ADAM33 is the most thoroughly investigated gene in this thesis, and SNPs in this gene have been associated with multiple clinical and immunological phenotypes in different study populations. We therefore hypothesize that the *ADAM33* gene may be a gene that affects lung function decline irrespective of disease, and that it modifies the severity of disease in established COPD as shown by lower lung function levels, faster FEV_1 decline, and more severe BHR and inflammation in COPD patients. Carriers of one or more *ADAM33* SNPs may develop different types of respiratory disease dependent on environmental factors and other genetic risk factors.

The *TGF-β1* 3'UTR SNP was associated with COPD development, whereas the Leu10Pro and C-509T SNPs in the same gene are associated with lower levels of FEV_1/VC in subjects with COPD from both the GRIP and Vlagtwedde/Vlaardingen population. This indicates again that it may well be that certain SNPs are only important in COPD development, whereas other SNPs are important in progression of the disease.

The importance of correct phenotyping has been shown in chapter 11 in which we could only replicate associations from the genetically isolated population in the general

population of Vlagtwedde/Vlaardingen when we compared COPD patients with similar disease severity. Even when using an objective measure like lung function, it is important to carefully select patient groups with similar disease severity when comparing them.

Definition of COPD

We have defined COPD according to the internationally used GOLD criteria [25]. A current debate is ongoing whether these definitions are correct since they are based on a fixed ratio of the post-bronchodilator FEV_1/FVC of 0.70 as a threshold. The choice for this threshold was a pragmatic one, based on the desire for simplicity, as perception of the complexity of spirometry and its interpretation has been one of the barriers to more widespread uptake of this clinical tool. A clinical diagnosis of COPD requires a history of chronic progressive symptoms, which are corroborated by abnormalities found on physical examination and a consideration of risk factors. However, objective evidence of airway obstruction determined by forced expiratory spirometry is the standard for demonstrating and quantifying airway obstruction [25]. Since the FEV_1/FVC ratio declines with age, using the fixed ratio to define COPD may over-diagnose COPD in older populations. Following the same rationale, the fixed ratio potentially under-diagnoses COPD in younger adults [26-28]. A recent American Thoracic Society and European Respiratory Society joint Task Force report recommends using a lower limit of normal (LLN) FEV_1/FVC as opposed to a fixed ratio of <0.70 to diagnose airway obstruction, in order to reduce false positive diagnoses of COPD as defined by the GOLD criteria. Yet, the GOLD criteria may prove to be valuable since in a longitudinal study subjects classified as "normal" using the LLN but abnormal when using the fixed ratio were more likely to die and to have a COPD-related hospitalization during follow-up. This suggests that a fixed FEV_1/FVC ratio of <0.70 may identify at-risk patients, even among older adults [29]. In the Vlagtwedde/Vlaardingen cohort, we have potentially overestimated the number of COPD cases by using the GOLD criteria. It would be worthwhile to take into account different definitions of COPD in future studies, and moreover to focus on phenotypes that are unquestionably valuable like levels of crude lung function and decline in lung function.

COPD occurs in different morphological subtypes. Some patients develop more airway obstruction, without emphysema, whereas others develop more small airways disease, or a combination of these two phenomena. The severity of emphysema varies widely even in patients with the same stage of COPD, and chronic bronchitis symptoms are equally distributed irrespective of emphysema severity [30]. Evidence of emphysema has been documented on CT scan images, even in subjects who have no airway obstruction as assessed by spirometry [31;32]. Abnormalities in FEV_1 have been associated with basal-predominant emphysema [33]. In addition to different COPD subtypes, there exist different morphologies between male and female COPD patients. As visible on CT-scans, males develop more regional and total emphysema and females more airways disease on all

stages of COPD severity [34;35]. These differences in phenotypes and between gender may have a genetic background as already demonstrated by Demeo *et al* [36]. The usage of CT scans should be incorporated in new studies on genetics of COPD.

Bronchial hyperresponsiveness (BHR)

BHR is defined as an exaggerated airway response to non-specific stimuli resulting in airway obstruction. BHR is a known risk factor for COPD development [37;38] and a majority of COPD patients demonstrates BHR [6;39]. The genetics of BHR is difficult to perform and is not very well described in existing literature in the context of the general population, and even less well in the context of COPD. The major problem in studying BHR is that it often co-occurs with other phenotypes like lower lung function, which makes it difficult to disentangle whether a SNP is associated with BHR or the co-occurring phenotype. In chapter 9 we therefore used multinomial regression to analyze whether SOD SNPs were associated with pure BHR, pure COPD, or a combination of the two phenotypes. These analyses showed that the pure phenotype BHR was associated with the *SOD-2* Ala16Val substitution and the *SOD-2* C7693T SNP with BHR with co-occurring COPD. This type of analyses is difficult to perform since large sample sizes are needed to obtain sufficiently large groups for comparisons.

In the GLUCOLD study, we determined that the *ADAM33* ST+5 SNP was associated with severity of BHR. Interestingly, *ADAM33* has originally been identified as a candidate gene for asthma, but the linkage reached the highest significance when including BHR in the asthma phenotype in the populations under study [1]. We did not find association of this gene with BHR in the general population (unpublished data). Our studies thus suggest that *ADAM33* is a gene that affects BHR severity in patients with existing COPD.

Markers of inflammation

In the GLUCOLD population, we had the unique opportunity to directly link genes to inflammation in sputum and bronchial biopsies. In the general population it would, apart from ethical aspects, not be feasible to perform biopsies and sputum induction on a large scale. On a population level, it would therefore be interesting to find systemic biomarkers in blood that reflect the inflammatory patterns in the lung. C reactive protein (CRP) is a marker of systemic inflammation and is associated with reduced FEV_1 and a predictor of accelerated decline in FEV_1 [40]. Moreover, CRP is a strong and independent predictor of future COPD outcomes in individuals with airway obstruction [41]. A study by Hersh *et al.* has shown a significant genetic influence on CRP levels in COPD with a heritability of 0.25 [42]. Genome-wide linkage analysis demonstrated no significant linkage of CRP to any chromosomal region, but revealed several potentially interesting chromosomal regions of borderline significance. A short tandem repeat marker near the *SFTPB* gene was significantly associated with CRP levels. In another study, surfactant protein D has recently

been identified as a serum marker for FEV_1 change and change in health status as reflected by CRQ scores [43]. It is yet unknown if genetic variation in *SFTPD* gene has an effect on serum levels of surfactant protein D.

Currently, more studies are ongoing to identify new serum biomarkers for COPD with promising results [44;45]. In future studies in the general population as well as in COPD patient cohorts, such serum biomarkers should be collected and used to determine the effect of genetic variation on levels of these markers.

SELECTION OF GENES AND POLYMORPHISMS

In our studies, we have investigated genes that had either previously been associated with COPD or phenotypes closely related to this disease or the genes had a biologically plausible role in COPD pathogenesis. The genotyped SNPs were consistently chosen based on data present in previous literature. This causes a deficiency in our studies, since these SNPs do not cover all the genetic variation in the gene, and hence we can miss associations or cannot pinpoint to the causal genetic variation. The way genes are analyzed nowadays is by using data from the Hapmap program, which has been introduced in 2003 and is updated regularly [46;47]. The recent update of the Hapmap database (Phase II) contains more than 3.1 million genetic variations, and maps approximately 1 SNP per 1000 basepairs in the genome [48]. The more precise the genomic map is the more precise genetic loci for disease can be identified. The Hapmap database not only shows the genetic variation, but also shows how strong the linkage disequilibrium (LD) is between SNPs. Many SNPs are inherited together in haplotype blocks and with a limited number of SNPs (so-called 'haplotype tagging SNPs') we can thus cover all genetic variation of a particular gene or chromosomal regions. In this thesis we have used the haplotype tagging SNP approach for *MMP9* and *decorin*, and we found no significant associations with these SNPs, indicating that these genes do not play a role in lung function decline and COPD development. However, we did find significant association with candidate SNPs for COPD selected based on previous literature (i.e. in *ADAM33*, *TGF-β1*, *MMP2*, *TIMP1*, *SFTPA*, *SFTPD*). We cannot be sure though that these SNPs are the true causative SNPs for the associated phenotype, or that they represent association of a nearby SNP in LD with the associated SNP. In case of the association of *ADAM33* it is known that there exists a strong LD within the gene, but the LD is also extended around the gene [49]. Therefore, it is possible that we are analyzing completely the wrong gene, even though *ADAM33* is one of the most reproduced genes in genetics of lung diseases.

In addition to the improvement in availability of genetic information, genotyping techniques have advanced and have become easier and in its applicability and are offered at lower costs nowadays, which makes it also economically feasible to study more SNPs in

practice. High throughput SNP genotyping has become available by DNA micro arrays from Affymetrix, bead array technology from Illumina and matrix assisted laser desorption/ionization-of flight mass spectrometry (MALDI-TOF). Such advances in genetic research have it made possible to perform whole genome association studies using panels of 300,000 or more SNPs across the genome to cover 90% or more of all genetic variation of the genome. The costs for such screens have dropped to a few hundred Euros per sample and are becoming more and more affordable for a larger number of research groups. Since large sample sizes are required for genome wide association studies, international collaborations are formed to pool phenotypic and genotypic data from affected patients and controls. In these pooled datasets, one should be aware of population stratification due to differences in genetic make-up between populations [50;51]. A very interesting future perspective is performing a genome wide association analysis for lung function level and decline in the Vlagtwedde/Vlaardingen cohort. In addition, this provides us with the unique opportunity to assess genome wide association for BHR and eosinophilia, both risk factors for development of COPD, and to assess gene-smoking interactions.

STATISTICAL ISSUES

As already mentioned in a previous paragraph, developments in genetic research have taken a fast track. Large scale genotyping possibilities make it possible to include larger sample sizes for genetic association studies. With the availability of large databases with genetic information, new techniques must be incorporated to adequately analyze the data.

False positive associations

The chance of finding associations due to chance becomes greater when multiple comparisons are made by testing association of many SNPs with different outcomes. The best test to prove that an association is true is replication of the effect in an independent cohort. The Vlagtwedde/Vlaardingen cohort is unique in its size and the time of follow-up, and therefore it is difficult to find a suitable replication cohort. The Doetinchem cohort is partly suitable, but the follow-up period is much shorter (approximately 5 years) and there are only 2 measurement of lung function, although data from a third cross-sectional survey in the Doetinchem study will follow soon. Nevertheless, these are low numbers of lung function measurements at surveys compared to the median of 7 we have from the Vlagtwedde/Vlaardingen participants. This makes the Doetinchem cohort less powerful to detect small genetic effects. However, if the associations are replicated in both cohorts, it is very likely a true effect. Since the Doetinchem cohort just recently became available as a replication cohort, we have tested only a few SNPs that were significant in our other

studies, for replication. Out of 5 SNPs that showed significant associations with lung function parameters in our previous studies, we could only replicate one significant association (*TIMP1* Phe124Phe). Since replication is increasingly becoming the golden standard in genetic epidemiology, we are indebted to continue with replication of significant associations in both cohorts. Lack of replication may be due to an initially false positive association, or may be the result of differences in study design, i.e. case definition, control selection, sample size, or population specific genetic or environmental specific modifiers [52-55]. Another strong warranty for successful replication is the need for large populations to ensure sufficient power to detect relatively small genetic effects [56].

Another strategy to deal with potential false positive associations is multiple testing corrections. Since SNPs in a gene are correlated and often phenotypes are related as well, a Bonferroni correction may be too rigid. Alternatively, in some of our studies, we used permutation testing and bootstrapping to generate a distribution of the beta-estimates under the null hypothesis. If the observed estimate occurs in less than 5% of the empiric cumulative distribution (p=0.05), one can assume that the observed beta-estimate is not found due to chance. In our studies, these types of methods confirmed the significant associations.

Haplotype analyses

Haplotypes are arrangements of alleles on the same chromosome inherited as a unit. In family based research, the precise combination of alleles in a haplotype can be determined using inheritance information of parent to child. In studies based on unrelated individuals, the parental origin of each allele is unknown and therefore the haplotypes for double heterozygous alleles must be estimated. There are several online tools available now to analyze haplotype associations with dichotomous outcomes, such as Merlin, Haploview, and Chaplin [57-62]. Most programs use similar modifications of the maximum likelihood method. Caution must be taken, however, to the proper use of these programs, since the slight differences in the way these programs estimate haplotype frequencies and differ in testing for associations may subsequently lead to different outcomes in different studies.

Haplotype analyses with quantitative traits was until recently not possible, but there are now available since a short period of time [63;64]. In these types of analyses, haplotype frequencies are estimated in the total population and an estimated chance per haplotype is established per subject in the dataset. These estimated chances can be used in the standard association analyses for quantitative traits. In our studies, we have not yet performed such analyses, but merely constructed phased multi-locus genotypes and tested for associations with quantitative phenotypes. In future studies the new techniques of using estimated haplotype frequencies should be incorporated.

Gene-environment interactions

Environmental factors may cause residual confounding and cloud genetic effects or obscure true associations. The environmental factor cigarette smoking plays a major role in COPD pathogenesis, and may in particular influence genetic factors like antioxidant genes. If subjects are not exposed to substances that induce oxidative stress, genetic variants will not be important in onset of the disease, under study compared to subjects with the same genetic make-up but *with* such exposures. A good example for such a phenomenon is the outcome of our own study in chapter 8. We observed that the effects of the *GSTM1* and *GSTT1* null alleles were particularly associated in male smokers. If we had not analyzed interaction of the genes with smoking we would not be able to identify subjects at risk.

Similarly, other environmental factors may cloud the genetic effects, or may obscure the populations at particular risk. Recently, Harber *et al.* showed that occupational exposure to fumes is a risk factor for lower annual lung function levels in early COPD [65]. This increased risk may be associated, and possibly partially be explained, by genetic susceptibility to the noxious effects of exposure to fumes. In our chapter on *GST*s we also addressed the effects of exogenous intake of antioxidant intake such as vitamin C. We unfortunately do not have information on dietary intake in the Vlagtwedde/Vlaardingen cohort. An interesting future perspective is that we have extensive data on diet via food questionnaires in the Doetinchem cohort. This makes this cohort suitable to assess interaction of genes and diet (and smoking) on the course of lung function prospectively. Another option is to perform stratified analyses or select study populations with similar environmental exposures [66].

Gene-gene interactions

SNPs in a gene may modify the action of SNPs in other genes that act in the same pathway. This phenomenon is known as epistasis. Such gene-gene interaction effects may enhance the risk for a particular phenotype in an additive or even multiplicative way, or may neutralize the negative effect of one SNP by the positive effect of another SNP [67]. Epistasis may play a major role in complex genetic diseases, since many genes are involved and it is therefore important to test for gene-gene interaction. However, large sample sizes with sufficient power are required to detect such gene-gene interactions, and appropriate sophisticated statistical methods, like the multifactor dimensionality reduction (MDR) method, are warranted to detect gene-gene interactions or gene-environment interactions [68-70]. Thus far, we have assessed gene-gene interactions between *decorin* and *TGF-β1* and additive effects of the combination of carrying *GSTM1* and *GSTT1* null alleles. These analyses did not provide significant results. However, we cannot exclude that the lack of association may have been caused by insufficient power.

FUNCTIONAL EFFECTS OF POLYMORPHISMS

The first step that needs to be made after a SNP has been proven to be associated with a certain phenotype in multiple independent studies is to analyze the functional effects of SNPs. The function of promoter SNPs is probably the most easy to determine as they may affect transcription levels of the gene. There are tools on the internet to analyze whether DNA sequences harbor transcription factor binding sites, and whether they are modified by SNPs or microsatellites in the promoter region. If this is theoretically the case, further experiments using for example luciferase assays may prove the effect in an *in vitro* assay. The final step is to assess whether *in vivo* such effects are observed as well, indicated in COPD for example by higher levels of the protein product in sputum. The *TGF-β1* C-509T SNP has proven to affect transcription levels of *TGF-β1* and also higher levels of TGF-β1 have been found in sputum from COPD patients. It is, however, still unclear if the increased TGF-β1 levels in sputum are related to SNPs in the gene.

Specific types of polymorphisms that often occur in promoter regions are microsatellites. These polymorphisms are repetitive sequences in the DNA and may affect transcription levels as suggested for example for the microsatellite polymorphisms in the *hemeoxygenase-1* gene in susceptibility for emphysema [71]. We have not studied microsatellites thus far in the Vlagtwedde/Vlaardingen cohort, but they may be studied in the future.

SNPs that are located in exons may change the amino acid sequence of a gene and can thereby modify the protein structure. This may alter the function of the protein directly, for example by altering the affinity of the protein to bind to its substrates. An amino acid change may also indirectly modify protein function, by altering the three dimensional folding of the protein. The advances in analyzing the functional effects of SNPs in the context of COPD are scarce. Some of the SNPs described in this thesis have been proven to have functional effects: *TGF-β1* Leu10Pro, *GSTP1* Ile105Val, *SFTPD* Met11Thr [72-78]. The same phenomenon applies for coding SNPs as for the promoter SNPs: studies on their functional relevance have been performed in *in vitro* systems, and the functional relevance has not been examined in humans with a direct link of SNP and effect, in COPD pathogenesis.

The function of non coding SNPs is difficult to comprehend, since they do not change the amino acid sequence and a direct effect on protein structure cannot be determined. However, a recent study in cardiovascular disease reported the first functional effect of a non coding polymorphism in the 3'UTR of a gene [79]. A polymorphism in the *angiotensin II type 1 receptor* (+1166 A/C) caused a decreased affinity of the RNA product with a specific microRNA (miRNA: miR-155). miRNAs are noncoding RNAs that silence gene expression by base-pairing with complementary sequences in the 3'UTR region of target RNAs. When the +1166 C-allele is present, base-pairing complementarity is interrupted, and the ability of

miR-155 to interact with the cis-regulatory site is decreased. As a result, miR-155 no longer attenuates translation as efficiently as with the A allele. The identification of such epi-genetic mechanisms is very complicated and the research is time consuming, but it shows that all genetic variations, and not just the easily comprehensible ones, deserve further investigation.

The third type of SNPs that are described in this thesis are the SNPs located in introns. Since introns are non coding gene regions, the effect of SNPs located in these regions is difficult to establish. Recently, an intron SNP in *ADAM33* has been described to influence expression of ADAM33 [80]. The allele that was associated with increased risk of asthma had a repressive effect on the promoter resulting in no detection of ADAM33 protein product. As a consequence, SNP BC+1 may have an important role in the modulation of *ADAM33* gene expression. We did not genotype this SNP in our population, and this should be performed in future studies on *ADAM33*.

Another suggested effect of intron SNPs is alternative splicing of genes, which is an efficient means of generating variation in protein function. Hull *et al*. suggested that phenotypic variation in splicing patterns is determined by the presence of SNPs within flanking introns or exons. Effects on splicing may represent an important mechanism by which SNPs influence gene function [81]. The intron SNPs that we have studied in this thesis (among others in *decorin* and *ADAM33*) do not possess functional effects, as far as we are aware of.

Finally, deletion polymorphisms may play a role in disease. The deletion alleles of *GSTM1* and *GSTT1* cause enzymatic deficiencies and therefore have an obvious effect on onset of disease. However, such large deletions do not occur very frequently and will not play a big role in complex genetic disorders.

USE OF GENETIC RISK FACTORS IN CLINICAL PRACTICE

It is of interest to know who is at risk for a disease by means of analyzing the genetic make-up in the future, which makes it possible to intervene at an early stage of disease. Another intriguing opportunity of knowing genetic risk factors is using specific medication targeted to solving deficits or excess expression of certain proteins due to genetic variation. One could think of supplements of antioxidant substances in case of a lack of antioxidant enzyme activity due to *GSTM1* deletion alleles, or supplying metalloprotease inhibitors in case of excess MMP expression. SNPs may also alter the responsiveness to medication. Coding SNPs in the β-adrenergic receptor have been shown to alter responsiveness to steroids in asthma and in COPD [82]. Such pharmacogenetic approaches may hopefully prove to be feasible in treatment of COPD patients [83].

FUTURE STUDIES

In the previous paragraphs several issues have been discussed that could be improved, or that are currently lacking, in genetic studies in COPD. In already existing cohorts, like the Vlagtwedde/Vlaardingen cohort for which data collection has stopped, we can only improve the method for analyzing the data and the choice of SNPs that are to be genotyped. In new to set-up studies, we should include, if financially possible, CT scans to disentangle the different subtypes of COPD. Currently lacking aspects in COPD genetics are a genome wide association of lung function level and decline in the general population, as well as mRNA profiles and proteomic profiles of subjects from the general population to identify new candidate genes for COPD. As previously stressed, it is important to identify genes that play a role in COPD development before the disease has become clinically manifest. Therefore such studies should be performed in a population with diverse, yet clearly defined, phenotypes and environmental exposures.

A diagram reflecting a potential study design is depicted in Figure 2. The diagram shows the starting point of collecting materials and data from COPD patients and subjects form the general population, such as exhaled breath condensate blood, DNA, sputum, lung and airway wall tissue, lung function data and CT-scans. These materials can be analyzed in whole genome association studies, mRNA profiling and proteomic analysis for novel genetic risk factors and biomarkers for COPD. Subsequently, the identified candidate genes and SNPs should ideally be replicated in multiple populations with similar power to detect the associations. In the next stage, the functionality of the replicated genes and SNPs should be assessed using both *in vitro* en *in vivo* models. In the last stage, it may be possible to develop genetic and proteomic risk profiles for subjects at risk to develop COPD or within patients with COPD, who are at risk to develop a more severe outcome of COPD.

An intriguing example of such a study, integrating data from genome wide association and mRNA data is the international multi-centre COPACETIC study (COPD Pathology: Addressing Critical gaps, Early Treatment & diagnosis and Innovative Concepts). This study will be performed in, among others, the University Medical Center Groningen. A genome wide association study will be performed on 4000 Dutch subjects with CT scans and lung function measurements (the NELSON study), and the 400 SNPs with the highest significance will be replicated in 5 populations, including the Vlagtwedde/Vlaardingen cohort. In addition, mRNA samples from subjects who are at high risk to develop COPD will be collected to discover molecular pathways. These pathways will be integrated with the significant SNPs from the genomic study. Using the information from these extended datasets, prediction models can be built to assess the capability to estimate the chance to develop COPD using both genomic and mRNA data. Moreover it allows assessing genetic profiles for specific types of COPD, i.e. (small) airway disease and emphysema.

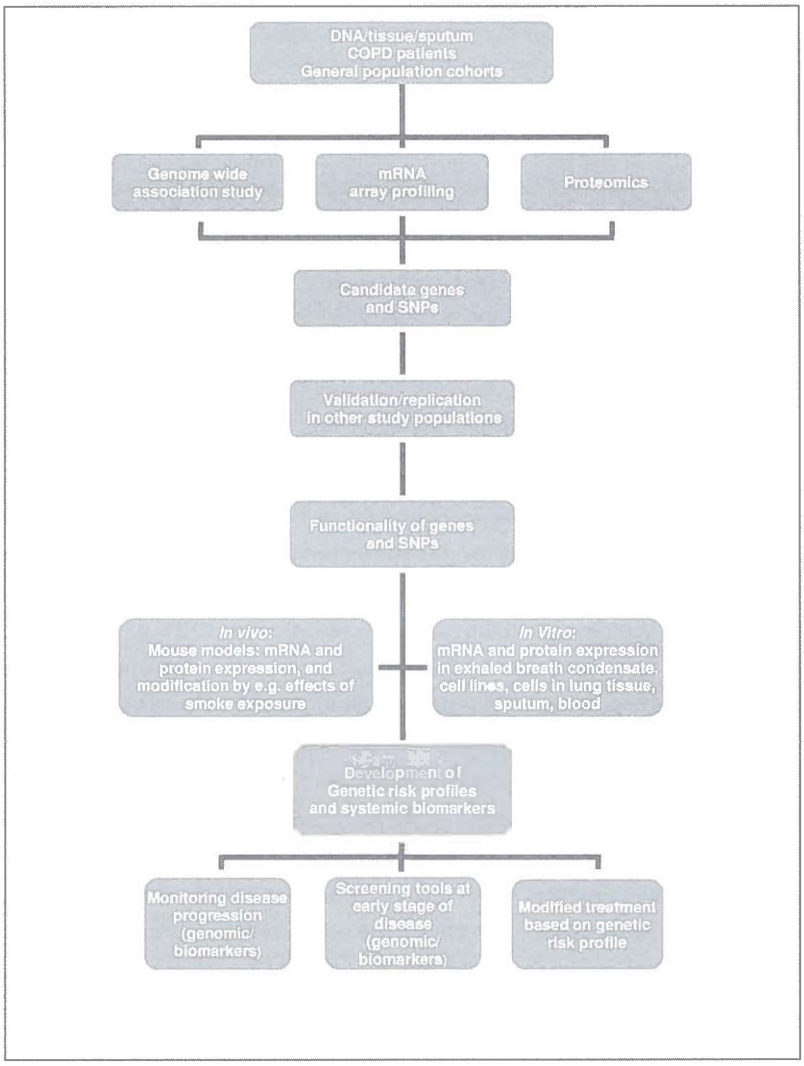


Figure 2: Flowchart of how future genetic studies for COPD may be set up, and to which purposes they can lead.

REFERENCES

1. Van Eerdewegh P, Little RD, Dupuis J, *et al.* Association of the ADAM33 gene with asthma and bronchial hyperresponsiveness. *Nature* 2002; 418:6896-430.

2. Howard TD, Postma DS, Jongepier H, *et al.* Association of a disintegrin and metalloprotease 33 (ADAM33) gene with asthma in ethnically diverse populations. *J Allergy Clin Immunol* 2003; 112:4-722.

3. Lee JH, Park HS, Park SW, *et al.* ADAM33 polymorphism: association with bronchial hyper-responsiveness in Korean asthmatics. *Clin Exp Allergy* 2004;34(6):860-5.

4. Werner M, Herbon N, Gohlke H, *et al.* Asthma is associated with single-nucleotide polymorphisms in ADAM33. *Clin Exp Allergy* 2004;34(1):26-31.

5. Raby BA, Silverman EK, Kwiatkowski DJ, *et al.* ADAM33 polymorphisms and phenotype associations in childhood asthma. *J Allergy Clin Immunol* 2004;113(6):1071-8.

6. Tashkin DP, Altose MD, Bleecker ER, *et al.* The lung health study: airway responsiveness to inhaled methacholine in smokers with mild to moderate airflow limitation. The Lung Health Study Research Group. *Am Rev Respir Dis* 1992;145(2 Pt 1):301-10.

7. Umland SP, Garlisi CG, Shah H, *et al.* Human ADAM33 Messenger RNA Expression Profile and Post-Transcriptional Regulation. *Am J Respir Cell Mol Biol* 2003; 29:5-582.

8. Garlisi CG, Zou J, Devito KE, *et al.* Human ADAM33: Protein maturation and localization. *Biochem Biophys Res Commun* 2003;301:1-43.

9. Powell RM, Wicks J, Holloway JW, *et al.* The splicing and fate of ADAM33 transcripts in primary human airways fibroblasts. Am J Respir Cell Mol Biol 2004;31:13-21.

10. Haitchi HM, Powell RM, Shaw TJ, *et al.* A Disintegrin and Metalloprotease 33 Expression in Asthmatic Airways and Human Embryonic Lungs. *Am J Respir Crit Care Med* 2005; 171(9):958-65.

11. Lee JY, Park SW, Chang HK, *et al.* A disintegrin and metalloproteinase 33 protein in patients with asthma: relevance to airflow limitation. *Am J Respir Crit Care Med* 2006;173(7):729-35.

12. Murray CJ, Lopez AD. Alternative projections of mortality and disability by cause 1990-2020: Global Burden of Disease Study. *Lancet* 1997;349(9064):1498-504.

13. Barnes PJ, Hansel TT. Prospects for new drugs for chronic obstructive pulmonary disease. *Lancet* 2004;364(9438):985-96.

14. He JQ, Ruan J, Connett JE, Anthonisen NR, *et al.* Antioxidant gene polymorphisms and susceptibility to a rapid decline in lung function in smokers. *Am J Respir Crit Care Med* 2002 Aug 1;166(3):323-8.

15. He JQ, Connett JE, Anthonisen NR, *et al.* Polymorphisms in the IL13, IL13RA1, and IL4RA genes and rate of decline in lung function in smokers. *Am J Respir Cell Mol Biol* 2003;28(3):379-85.

16. He JQ, Connett JE, Anthonisen NR, *et al.* Glutathione S-transferase variants and their interaction with smoking on lung function. *Am J Respir Crit Care Med* 2004 Aug 15;170(4):388-94.

17. Friedman GD, Klatsky AL, Siegelaub AB. Lung function and risk of myocardial infarction and sudden cardiac death. *N Engl J Med* 1976;294(20):1071-5.

18. Lange P, Nyboe J, Jensen G, *et al.* Ventilatory function impairment and risk of cardiovascular death and of fatal or non-fatal myocardial infarction. *Eur Respir J* 1991;4(9):1080-7.

19. Marcus EB, Curb JD, MacLean CJ, *et al.* Pulmonary function as a predictor of coronary heart disease. *Am J Epidemiol* 1989;129(1):97-104.

20. Weiss ST, Segal MR, Sparrow D, *et al.* Relation of FEV1 and peripheral blood leukocyte count to total mortality. The Normative Aging Study. *Am J Epidemiol* 1995;142(5):493-8.

21. Gilliland FD, Gauderman WJ, Vora H, *et al.* Effects of glutathione-S-transferase M1, T1, and P1 on childhood lung function growth. *Am J Respir Crit Care Med* 2002;166(5):710-6.

22. Schedel M, Depner M, Schoen C, *et al.* The role of polymorphisms in ADAM33, a disintegrin and metalloprotease 33, in childhood asthma and lung function in two German populations. *Respir Res* 2006;7:91.

23. Simpson A, Maniatis N, Jury F, *et al.* Polymorphisms in A Disintegrin and Metalloprotease 33 Predict Impaired Early-Life Lung Function. *Am J Respir Crit Care Med* 2005;172(1):55-60.

24. DeMeo DL, Campbell EJ, Barker AF, *et al.* IL10 Polymorphisms are Associated with Airflow Obstruction in Severe alpha 1-antitrypsin Deficiency. *Am J Respir Cell Mol Biol* 2008;38(1):114-120.

25. From the Global Strategy for the Diagnosis, Management and Prevention of COPD, Global Initiative for Chronic Obstructive Lung Disease (GOLD) 2006. Available from: http://www.goldcopd.org. 2007.

26. Aggarwal AN, Gupta D, Behera D, *et al.* Comparison of fixed percentage method and lower confidence limits for defining limits of normality for interpretation of spirometry. *Respir Care* 2006;51(7):737-43.

27. Roberts SD, Farber MO, Knox KS, *et al.* FEV1/FVC ratio of 70% misclassifies patients with obstruction at the extremes of age. *Chest* 2006;130(1):200-6.

28. Hardie JA, Buist AS, Vollmer WM, *et al.* Risk of over-diagnosis of COPD in asymptomatic elderly never-smokers. *Eur Respir J* 2002;20(5):1117-22.

29. Mannino DM, Sonia BA, Vollmer WM. Chronic obstructive pulmonary disease in the older adult: what defines abnormal lung function? *Thorax* 2007;62(3):237-41.

30. Makita H, Nasuhara Y, Nagai K, *et al.* Characterisation of phenotypes based on severity of emphysema in chronic obstructive pulmonary disease. *Thorax* 2007;62(11):932-7.

31. Gurney JW, Jones KK, Robbins RA, *et al.* Regional distribution of emphysema: correlation of high-resolution CT with pulmonary function tests in unselected smokers. *Radiology* 1992;183(2):457-63.

32. Klein JS, Gamsu G, Webb WR, *et al.* High-resolution CT diagnosis of emphysema in symptomatic patients with normal chest radiographs and isolated low diffusing capacity. *Radiology* 1992;182(3):817-21.

33. Holme J, Stockley RA. Radiologic and clinical features of COPD patients with discordant pulmonary physiology: lessons from alpha1-antitrypsin deficiency. *Chest* 2007;132(3):909-15.

34. Dransfield MT, Washko GR, Foreman MG, *et al.* Gender differences in the severity of CT emphysema in COPD. *Chest* 2007;132(2):464-70.

35. Martinez FJ, Curtis JL, Sciurba F, *et al.* Gender Differences in Severe Pulmonary Emphysema. *Am J Respir Crit Care Med* 2007; 176(3):243-52.

36. DeMeo DL, Hersh CP, Hoffman EA, *et al.* Genetic Determinants of Emphysema Distribution in the National Emphysema Treatment Trial. *Am J Respir Crit Care Med* 2007;176(1):42-8..

37. Postma DS, Boezen HM. Rationale for the Dutch hypothesis. Allergy and airway hyperresponsiveness as genetic factors and their interaction with environment in the development of asthma and COPD. *Chest* 2004;126(2 Suppl):96S-104S.

38. Brutsche MH, Downs SH, Schindler C, *et al.* Bronchial hyperresponsiveness and the development of asthma and COPD in asymptomatic individuals: SAPALDIA cohort study. *Thorax* 2006;61(8):671-7.

39. Tashkin DP, Altose MD, Connett JE, *et al.* Methacholine reactivity predicts changes in lung function over time in smokers with early chronic obstructive pulmonary disease. The Lung Health Study Research Group. *Am J Respir Crit Care Med* 1996;153(6 Pt 1):1802-11.

40. Fogarty AW, Jones S, Britton JR, *et al.* Systemic inflammation and decline in lung function in a general population: a prospective study. *Thorax* 2007;62(6):515-20.

41. Dahl M, Vestbo J, Lange P, *et al* C-reactive protein as a predictor of prognosis in chronic obstructive pulmonary disease. *Am J Respir Crit Care Med* 2007;175(3):250-5.

42. Hersh CP, Miller DT, Kwiatkowski DJ, *et al.* Genetic determinants of C-reactive protein in COPD. *Eur Respir J* 2006;28(6):1156-62.

43. Sin DD, Leung R, Gan WQ, *et al.* Circulating surfactant protein D as a potential lung-specific biomarker of health outcomes in COPD: a pilot study. *BMC Pulm Med* 2007;7(1):13.

44. Bandow J, Sepulveda J, Baker J, *et al.* Proteomic analysis of plasma from COPD patients and smoking control subjects . *Conference of Eur Resp Soc* 2007,abstract 38s.

45. Kilty I, Walsh R, Bandow J, *et al.* Application of omics technologies to the identification of biomarkers fro COPD. *Conference of Eur Resp Soc* 2007, 212s.

46. The International HapMap Project. *Nature* 2003;426(6968):789-96.

47. Altshuler D, Brooks LD, Chakravarti A, *et al.* A haplotype map of the human genome. *Nature* 2005;437(7063):1299-320.

48. Frazer KA, Ballinger DG, Cox DR, *et al.* A second generation human haplotype map of over 3.1 million SNPs. *Nature* 2007;449(7164):851-61.

49. Wjst M. Public data mining shows extended linkage disequilibrium around ADAM33. *Allergy* 2007;62(4):444-6.

50. Keinan A, Mullikin JC, Patterson N, *et al.* Measurement of the human allele frequency spectrum demonstrates greater genetic drift in East Asians than in Europeans. *Nat Genet* 2007;39(10):1251-5.

51. Ioannidis JP, Ntzani EE, Trikalinos TA. 'Racial' differences in genetic effects for complex diseases. *Nat Genet* 2004;36(12):1312-8.

52. Ioannidis JP, Ntzani EE, Trikalinos TA, *et al*. Replication validity of genetic association studies. *Nat Genet* 2001;29(3):306-9.

53. Cardon LR, Palmer LJ. Population stratification and spurious allelic association. *Lancet* 2003;361(9357):598-604.

54. Cardon LR, Bell JI. Association study designs for complex diseases. *Nat Rev Genet* 2001;2(2):91-9.

55. Cordell HJ, Clayton DG. Genetic association studies. *Lancet* 2005;366(9491):1121-31.

56. Ioannidis JP, Trikalinos TA, Ntzani EE, *et al.* Genetic associations in large versus small studies: an empirical assessment. *Lancet* 2003;361(9357):567-71.

57. Abecasis GR, Cherny SS, Cookson WO, *et al*. Merlin--rapid analysis of dense genetic maps using sparse gene flow trees. *Nat Genet* 2002;30(1):97-101.

58. Stephens M, Smith NJ, Donnelly P. A new statistical method for haplotype reconstruction from population data. *Am J Hum Genet* 2001;68(4):978-89.

59. Falush D, Stephens M, Pritchard JK. Inference of population structure using multilocus genotype data: linked loci and correlated allele frequencies. *Genetics* 2003;164(4):1567-87.

60. Scheet P, Stephens M. A fast and flexible statistical model for large-scale population genotype data: applications to inferring missing genotypes and haplotypic phase. *Am J Hum Genet* 2006;78(4):629-44.

61. Case-Control Haplotype Interference (CHAPLIN), version 1.2. 2006.

62. Epstein MP, Satten GA. Inference on haplotype effects in case-control studies using unphased genotype data. *Am J Hum Genet* 2003;73(6):1316-29.

63. Sham PC, Rijsdijk FV, Knight J, *et al.* Haplotype association analysis of discrete and continuous traits using mixture of regression models. *Behav Genet* 2004;34(2):207-14.

64. Tzeng JY, Wang CH, Kao JT, *et al.* Regression-based association analysis with clustered haplotypes through use of genotypes. *Am J Hum Genet* 2006;78(2):231-42.

65. Harber PI, Tashkin DP, Simmons M, *et al.* Effect of Occupational Exposures on Decline of Lung Function in Early COPD. *Am J Respir Crit Care Med* 2007; 176(10):994-1000.

66. Boks MP, Schipper M, Schubart CD, *et al.* Investigating gene environment interaction in complex diseases: increasing power by selective sampling for environmental exposure. *Int J Epidemiol* 2007;36(6):1363-9.

67. Cordell HJ. Epistasis: what it means, what it doesn't mean, and statistical methods to detect it in humans. *Hum Mol Genet* 2002;11(20):2463-8.

68. Lee SY, Chung Y, Elston RC, *et al.* Log-linear model-based multifactor dimensionality reduction method to detect gene gene interactions. *Bioinformatics* 2007;23(19):2589-95.

69. Ritchie MD, Hahn LW, Roodi N, *et al.* Multifactor-dimensionality reduction reveals high-order interactions among estrogen-metabolism genes in sporadic breast cancer. *Am J Hum Genet* 2001;69(1):138-47.

70. Ritchie MD, Hahn LW, Moore JH. Power of multifactor dimensionality reduction for detecting gene-gene interactions in the presence of genotyping error, missing data, phenocopy, and genetic heterogeneity. *Genet Epidemiol* 2003;24(2):150-7.

71. Yamada N, Yamaya M, Okinaga S, *et al.* Microsatellite polymorphism in the heme oxygenase-1 gene promoter is associated with susceptibility to emphysema. *Am J Hum Genet* 2000;66(1):187-95.

72. Silverman ES, Palmer LJ, Subramaniam V, *et al.* Transforming Growth Factor-beta1 Promoter Polymorphism C-509T Is Associated with Asthma. *Am J Respir Crit Care Med* 2004;169(2):214-9.

73. Hobbs K, Negri J, Klinnert M, *et al.* Interleukin-10 and transforming growth factor-beta promoter polymorphisms in allergies and asthma. *Am J Respir Crit Care Med* 1998;158(6):1958-62.

74. Grainger DJ, Heathcote K, Chiano M, *et al.* Genetic control of the circulating concentration of transforming growth factor type beta1. *Hum Mol Genet* 1999 Jan;8(1):93-7.

75. Awad MR, El Gamel A, Hasleton P, *et al.* Genotypic variation in the transforming growth factor-beta1 gene: association with transforming growth factor-beta1 production, fibrotic lung disease, and graft fibrosis after lung transplantation. *Transplantation* 1998;66(8):1014-20.

76. Sundberg K, Johansson AS, Stenberg G, *et al*. Differences in the catalytic efficiencies of allelic variants of glutathione transferase P1-1 towards carcinogenic diol epoxides of polycyclic aromatic hydrocarbons. *Carcinogenesis* 1998;19(3):433-6.

77. Heidinger K, Konig IR, Bohnert A, *et al*. Polymorphisms in the human surfactant protein-D (SFTPD) gene: strong evidence that serum levels of surfactant protein-D (SP-D) are genetically influenced. *Immunogenetics* 2005;57(1-2)1-7.

78. Leth-Larsen R, Garred P, Jensenius H, *et al*. A common polymorphism in the SFTPD gene influences assembly, function, and concentration of surfactant protein D. *J Immunol* 2005;174(3):1532-8.

79. Martin MM, Buckenberger JA, Jiang J, *et al*. The human angiotensin II type 1 receptor +1166 A/C polymorphism attenuates microrna-155 binding. *J Biol Chem* 2007;282(33):24262-9.

80. Del Mastro RG, Turenne L, Giese H, Keith TP, *et al*. Mechanistic role of a disease-associated genetic variant within the ADAM33 asthma susceptibility gene. *BMC Med Genet* 2007;8:46.

81. Hull J, Campino S, Rowlands K, *et al*. Identification of common genetic variation that modulates alternative splicing. *PLoS Genet* 2007;3(6):e99.

82. Yang IA, Ng T, Molenaar P, Fong KM. Beta2-adrenoceptor polymorphisms and obstructive airway diseases: important issues of study design. *Clin Exp Pharmacol Physiol* 2007;34(10):1029-36.

83. Pettipher R, Cardon LR. The application of genetics to the discovery of better medicines. *Pharmacogenomics* 2002;3(2):257-63.

Nederlandse Samenvatting

Summary in Dutch

Inleiding

Bij vrijwel ieder mens neemt de longfunctie met het ouder worden af, gemiddeld met 30 ml per jaar. Mensen die roken hebben vaak een versnelde achteruitgang van de longfunctie wat aanleiding kan geven tot klachten van de luchtwegen en longziekte. Een van de ziekten die kan ontstaan door roken is Chronic Obstructive Pulmonary Disease (COPD). COPD is de verzamelnaam voor chronische bronchitis en longemfyseem, aandoeningen die worden gekarakteriseerd door luchtwegobstructie, luchtwegontsteking en afbraak van longweefsel. Bij chronische bronchitis zijn er meer slijmproducerende cellen, waardoor er meer slijm in de luchtwegen komt en mensen benauwd worden en moeten hoesten. Bij emfyseem vindt afbraak plaats van de longblaasjes, zodat er minder zuurstof uit de ingeademde lucht kan worden opgenomen door het lichaam. Dit leidt tot kortademigheid en in ernstige gevallen van COPD moeten patiënten extra zuurstof toegediend krijgen. Veel COPD patiënten hebben zowel kenmerken van chronische bronchitis als van longemfyseem. Ongeveer 10% van de Nederlandse bevolking leidt aan COPD. De ziekte is chronisch en er zijn tot op heden nog geen goede medicijnen om de ziekte tot stilstand te kunnen brengen.

Karakterisering van Chronic Obstructive Pulmonary Disease

Een groot deel van de diagnose van COPD wordt gebaseerd op basis van longfunctietesten (spirometrie). Bij deze testen worden onder meer het totale longvolume en de maximale hoeveelheid uitgeademde lucht in 1 seconde gemeten, respectievelijk de vitale capaciteit (VC) en het Forced Expiratory Volume in 1 seconde (FEV_1). De inhoud van de longen kan ook geschat worden op basis van wat normaal is gegeven de leeftijd, lengte en het geslacht van een individu. Hoe verder de geblazen waarden afwijken van deze normaal-waarden, hoe ernstiger de COPD. Een indeling voor de ernst van COPD die wereldwijd veel gebruikt wordt is de zogenaamde GOLD indeling (Global Initiative for Obstructive Lung Diseases). Er bestaan 4 stadia, gebaseerd op de waarde van de FEV_1 en de VC (tabel 1).

Tabel 1: Classificatie van COPD gebaseerd op longfunctiemetingen

GOLD stadium	Longfunctie
I: Mild COPD	• $FEV_1/FVC < 70\%$ • $FEV_1 \geq 80\%$ voorspeld
II: Matig COPD	• $FEV_1/FVC < 70\%$ • $50\% \leq FEV_1 < 80\%$ voorspeld
III: Ernstig COPD	• $FEV_1/FVC < 70\%$ • $30\% \leq FEV_1 < 50\%$ voorspeld
IV: Zeer ernstig COPD	• $FEV_1/FVC < 70\%$ • $FEV_1 < 30\%$ voorspeld of $FEV_1 < 50\%$ voorspeld plus chronisch respiratoir falen

Risicofactoren voor COPD

Roken is vaak de belangrijkste oorzaak van COPD, hoewel niet iedere roker COPD ontwikkelt. In het bevolkingsonderzoek van Vlagtwedde/Vlaardingen, waarin mensen gedurende 30 jaar zijn vervolgd, bleken er naast roken nog een aantal andere risicofactoren te bestaan voor COPD. Mensen die hyperreactief zijn hebben bij inademen van prikkels als kou en mist luchtwegvernauwing en klachten als hoesten en benauwdheid. Deze mensen hebben een snellere achteruitgang van de longfunctie. Daarnaast hebben mensen met een hoger aantal eosinofiele cellen in het bloed en een verhoogd totaal serum IgE een toegenomen achteruitgang.

Uit de literatuur lijkt de hoogte van de longfunctie ook genetisch bepaald te zijn. Bij eeneiige tweelingen die roken blijkt de afname van de longfunctie vergelijkbaar te zijn, maar dat is minder het geval bij twee-eiige tweelingen die roken. Bovendien komt in families van COPD patiënten vaker COPD voor. Dat wijst erop dat er een erfelijke component aanwezig is. Vaak wordt genetisch onderzoek, om genen op te sporen die betrokken zijn bij een ziekte, verricht in families. Dat is bij COPD echter moeilijk omdat deze longaandoening meestal pas op middelbare leeftijd tot uiting komt, wanneer de ouders van de persoon met COPD vaak al zijn overleden.

Genetica

De genetische informatie van ieder organisme, ook wel het genoom genoemd, ligt opgeslagen in de chromosomen. Het humane genoom bestaat uit 23 paren van chromosomen, waarvan 1 paar de geslachtschromosomen zijn (X en Y). Chromosomen zijn opgebouwd uit DNA. Het DNA molecuul bestaat uit twee koordachtige strengen, lijkend op een ladder, die om elkaar heen zijn gedraaid. Deze structuur wordt ook wel de dubbele helix genoemd. DNA is opgebouwd uit 4 verschillende nucleotiden: Adenine, Cytosine, Thymine en Guanine, afgekort tot A, C, T, G. Deze komen altijd in paren voor op tegenovergestelde strengen DNA: **A+T** en **C+G**.
Een enkele streng van DNA bestaat dus uit deze 4 letters:
ATGCTCGAATAAATGATTTGA
Deze letters maken woorden van drie nucleotiden die ieder coderen voor een aminozuur:
ATG CTC GAA TAA ATG ATT TGA
Een combinatie van deze woorden (aminozuren) vormen zinnen, de genen:
<ATG CTC GAA TAA> <ATG ATT TGA>
Aminozuren zijn de bouwstenen van eiwitten, en genen bevatten dus de informatie hoe eiwitten geproduceerd moeten worden. Eiwitten zijn nodig voor de structuur, functie, en regulering van de cellen, weefsels en organen van het lichaam.

Een mutatie of polymorfisme is een verandering in het DNA in de chromosomen. Polymorfismen zijn veel voorkomende veranderingen in de volgorde van het DNA, die bij tenminste 1% van de algemene bevolking voorkomen. Een zogeheten mutatie komt in minder dan 1% van de algemene bevolking voor. De meeste variatie in DNA in neutraal, dat wil zeggen dat het geen negatieve of positieve consequenties heeft voor de functie van eiwitten, cellen of organen. Het kan echter voorkomen dat een verandering in een gen resulteert in een verandering van een eiwit, zodat dit een andere werking heeft of niet meer functioneert. Dergelijke variaties kunnen bijdragen aan de ontwikkeling van ziekte, de ernst van ziekte, of hoe het lichaam reageert op medicatie.

Van elk chromosoom hebben we er twee: 1 overgeërfd van de vader en 1 van de moeder. Het is dus mogelijk dat we van een polymorfisme op een chromosoom de normale variant hebben (het 'wild type'), en op de andere de mutatie. Dit noemen we heterozygoot. Als we op beide chromosomen de mutante vorm dragen, noemen we dit een homozygote variant.

DOEL VAN HET ONDERZOEK IN DIT PROEFSCHRIFT

Wij willen in het huidige onderzoek vaststellen of genetisch factoren het longfunctie-beloop tijdens het leven bepalen. Dit kan uiteindelijk tot COPD aanleiding geven. Wij hebben de unieke mogelijkheid om DNA te analyseren van deelnemers aan het Vlagtwedde/Vlaardingen onderzoek die 25 jaar vervolgd zijn en waarbij gegevens zijn verzameld zoals longfunctiemetingen, vragenlijsten met betrekking tot luchtwegklachten en rookgewoonten. Dit maakt dat wij grote groepen personen kunnen vergelijken met snelle of langzame achteruitgang in de longfunctie in relatie tot genetische en reeds bekende risicofactoren. We hebben onderzocht of polymorfismen in genen relevant zijn bij het ontstaan van COPD en de ernst van COPD. Hiertoe hebben we naast het Vlagtwedde/Vlaardingen cohort nog 3 groepen COPD patiënten onderzocht. In tabel 1 staan de onderzoekspopulaties weergegeven met hun belangrijkste kenmerken.

Tabel 1: Studie populaties beschreven in dit proefschrift

	Vlagtwedde/ Vlaardingen	Doetinchem	GLUCOLD	LTX	GRIP
Selectie criteria	Algemene bevolking	Algemene bevolking	COPD patiënten GOLD stadium ≥ II	COPD patiënten GOLD stadium ≥ III die een LTX hebben ondergaan	COPD patiënten GOLD stadium ≥ I van een genetisch isolaat
Aantal	1390	1152	114	79	106
Type data	Longitudinale longfunctie data, BHR, bloed eosino-filie en klachten gedurende 25 jaar	Longitudinale longfunctie data en klachten gedurende 5 jaar	Cross-sectionele data van inflammatie in sputum, bloed en bronchiale biopten, BHR, longfunctie	Longitudinale long - functiedata verkregen voor LTX van patiënten met en zonder α1-anti-trypsine deficiëntie	Cross-sectionele longfunctie data in een genetisch geïsoleerde populatie met genealogie data

Afkortingen: GLUCOLD Groningen Leiden Universities Corticosteroids in Obstructive Lung Disease; LTX long transplantatie; GRIP Genetic Research in Isolated Populations; COPD Chronic Obstructive Pulmonary Disease; GOLD Global initiative for Obstructive Lung Diseases; BHR Bronchiale Hyperreactiviteit

GENETISCHE EPIDEMIOLOGIE VAN LONGFUNCTIE

Er bestaan verschillende methoden om de genetische achtergrond van een ziekte te bestuderen. *Hoofdstuk 2* van dit proefschrift beschrijft een aantal veel gebruikte technieken in de genetica van longziekten. In *familie- en tweelingonderzoek* wordt onderzocht of er clustering is van een bepaalde ziekte in families of tussen tweelingen. Zo blijkt de hoogte van de longfunctie meer gelijk te zijn in een-eiïge tweelingen (die 100% gelijk DNA hebben) dan in twee-eiïge tweelingen (die voor 50% hetzelfde DNA hebben). In *'genomewide linkage analysis'* wordt bepaald welke delen van de chromosomen betrokken zijn bij het ontstaan van de ziekte. In de chromosomen bevinden zich kleine stukjes DNA die als 'marker' kunnen dienen. Als meerdere familieleden met ziekte steeds dezelfde marker hebben overgeërfd is dit een indicatie dat in de buurt van die marker zich een ziekteveroorzakend gen bevindt. In het onderzoek naar de erfelijkheid van longfunctie zijn al een aantal chromosomale gebieden gevonden die te maken hebben met de hoogte van de longfunctie, maar de causale genen zijn nog niet geïdentificeerd.

De meeste studies in genetische epidemiologie worden uitgevoerd in een *'case-control'* opzet. Hierbij worden genen onderzocht die op basis van hun biologische functie een rol zouden kunnen spelen in het ontstaan van ziekte ('kandidaat-genen') . Er wordt dan onderzocht of een bepaalde variant van een gen vaker voorkomt bij mensen met de ziekte dan bij mensen zonder de ziekte. Het is vaak moeilijk om exact dezelfde resultaten van dergelijke studies in andere populaties te vinden, omdat de ziekte verschillend gedefinieerd wordt, er etnische verschillen zijn, of omdat de onderzoekspopulaties te klein zijn.

In *longitudinale cohort studies* van de algemene bevolking is het mogelijk om genetische effecten op te sporen die betrokken zijn bij een verlaagde longfunctie voordat klinisch relevante ziekte is ontstaan. Helaas zijn dergelijke studies erg duur omdat veel personen moeten worden onderzocht en de studies lang duren.

Een *genetisch geïsoleerde populatie* is gesticht door een beperkt aantal voorvaders, en waarbij bovendien weinig uitwisseling is geweest met andere populaties in de loop der tijd. Hierdoor is de genetische diversiteit beperkt en kunnen bepaalde ziekteveroorzakende mutaties vaker voorkomen dan in een genetisch diverse populatie. Dit maakt genetisch geïsoleerde populaties geschikt om in een relatief kleine opzet toch ziektegeassocieerde genen op te sporen.

DE ROL VAN POLYMORFISMEN IN *ADAM33* BIJ DALING VAN LONGFUNCTIE IN DE ALGEMENE BEVOLKING EN ERNST VAN ZIEKTE BINNEN COPD PATIËNTEN

In 2003 is een gen gevonden via 'genome wide linkage analyse' dat een sterke associatie had met astma. Uit nader onderzoek in Groningen, bleek dat dit gen, genaamd *ADAM33* (*A Disintegrin And Metalloprotease 33*) geassocieerd is met een versnelde longfunctiedaling in astmapatiënten. Naar aanleiding van dit resultaat veronderstelden we dat dit gen ook betrokken kon zijn bij versnelde achteruitgang van longfunctie in de algemene bevolking. In *hoofdstuk 3*, hebben we de rol onderzocht van 8 polymorfismen in *ADAM33* in de Vlagtwedde/Vlaardingen bevolking. De daling in longfunctie voor de gehele populatie was gemiddeld 19.2 ml/jaar. Drie van de onderzochte polymorfismen bleken geassocieerd te zijn met versnelde achteruitgang van longfunctie (namelijk Q-1, S1 en S2) ; dragers van 1 bepaald polymorfisme (Q-1) hadden een 9.6 ml snellere daling van longfunctie per jaar dan dragers van de wild type variant. Ook bleken 4 polymorfismen (F+1, S1, S2 en T2) vaker voor te komen bij personen met COPD dan bij personen zonder COPD. Uit dit onderzoek blijkt *ADAM33* niet alleen een rol te spelen in de ontwikkeling van astma, maar ook in het verloop van longfunctie in de algemene bevolking en de ontwikkeling van COPD.

Vervolgens hebben we onderzocht of genetische variatie in *ADAM33* ook geassocieerd is met de ernst van COPD binnen patiënten met COPD (*hoofdstuk 4*). We hebben 111 patiënten met een mild tot matig ernstige vorm van COPD uit de GLUCOLD studie gegenotypeerd met 8 SNPs van *ADAM33*. Met lineaire regressiemodellen hebben we gekeken of SNPs geassocieerd waren met bronchiale hyperreactiviteit en ontstekingscellen in sputum en weefselbiopten van de luchtwegwand. We vonden een associatie van 1 SNP (ST+5) met zowel minder ernstige hyperreactiviteit, minder cellen in sputum en minder CD8+ cellen (een bepaald type immuuncel) in biopten. Ook waren twee andere SNPs geassocieerd met minder ontstekingscellen in biopten. Deze studie laat zien

dat SNPs in *ADAM33* ook betrokken zijn bij de ernst van COPD en het ontstekingsproces dat betrokken is bij het ontstaan van COPD.

α1-Antitrypsinedeficiëntie is een aangeboren afwijking die kan leiden tot de ontwikkeling van COPD op jonge leeftijd zonder dat de patiënt gerookt hoeft te hebben. Mensen met α1-antitrypsine deficiëntie hebben vanaf de geboorte een tekort aan het eiwit α1-antitrypsine, dat nodig is voor de bescherming van de longen tegen afbraak van longweefsel.

We hebben de rol van *ADAM33* bestudeerd in 79 patiënten met ernstig COPD, met en zonder α1-antitrypsinedeficiëntie, die een longtransplantatie hebben moeten ondergaan vanwege de ernst van hun COPD (*hoofdstuk 5*). We hebben longitudinale data van 79 patiënten teruggevraagd en uiteindelijk longfunctiegegevens van gemiddeld 5 jaar voor hun transplantatie verkregen (6 longfunctiemetingen, 42 personen met α1-antitrypsinedeficiëntie). Van deze patiënten hebben we DNA geïsoleerd uit bevroren longweefsel. We hebben SNPs in *ADAM33* gegenotypeerd in deze patiënten en geanalyseerd of er een associatie was met daling in FEV_1. We stelden vast dat 2 SNPs (ST+5 en T_2) geassocieerd waren met extra daling in FEV_1, onafhankelijk van α1-antitrypsinedeficiëntie. Deze resultaten laten voor het eerst zien dat genetische factoren anders dan α1-antitrypsinedeficiëntie, en onafhankelijk van het effect van α1-antitrypsinedeficiëntie, kunnen bijdragen aan versnelde longfunctiedaling bij COPD patiënten. Nader onderzoek naar die genetische factoren zou kunnen resulteren in het vinden van aangrijpingspunten voor andere dan tot nu toe in Nederland gebruikelijke behandelingen voor patiënten met α1-antitrypsinedeficiëntie.
Omdat de functie van *ADAM33* nog niet bekend is, en ook niet bekend is welke cellen *ADAM33* tot expressie brengen, hebben we de expressie van *ADAM33* bepaald in longweefsel van deze patiënten met behulp van immunohistochemie (*hoofdstuk 5*). Hierbij wordt longweefsel geïncubeerd met een eiwit dat specifiek ADAM33 herkent (een antilichaam). Dit antilichaam wordt vervolgens voorzien van een label, waardoor onder de microscoop zichtbaar wordt waar het eiwit (en dus ADAM33) zich bevindt. Met deze techniek hebben we sterke ADAM33 expressie in endotheel, spierweefsel, mesenchymale cellen en epitheel gevonden. Er waren geen verschillen in expressie tussen patiënten met en zonder α1-antitrypsinedeficiëntie, of tussen COPD en controle longweefsel. Aangezien de mate en locatie van expressie gelijk was tussen de patiënten konden we niet analyseren of de expressie verschillend was tussen patiënten met verschillende genetische varianten.

GENEN BETROKKEN BIJ AFBRAAK EN HERSTEL VAN LONGWEEFSEL IN COPD

COPD is gekarakteriseerd door luchtwegobstructie, chronische inflammatie van de luchtwegen en afbraak van longweefsel. Matrix Metalloproteases (MMPs) zijn enzymen die eiwitten kunnen knippen en betrokken zijn bij de normaal voorkomende reparatie van weefsel. Uit de literatuur is bekend dat deze MMPs verhoogd aanwezig zijn in de longen van emfyseempatiënten en betrokken zijn bij de afbraak van longweefsel in deze patiënten. Verder is gebleken dat SNPs in een aantal van deze *MMP*s geassocieerd zijn met daling in longfunctie. De Tissue Inhibitors van Metalloproteases (TIMPs) zijn de natuurlijke remmers van MMPs en kunnen de afbraak van longweefsel tegengaan. Eén van de heersende hypotheses over de ontwikkeling van COPD is dat de balans tussen MMPs en TIMPs niet goed gereguleerd is in COPD. Dit zou mogelijk een genetische achtergrond kunnen hebben. In *hoofdstuk 6* hebben we onderzocht of SNPs in *MMP1*, *MMP2*, *MMP9*, *MMP12* en *TIMP1* geassocieerd waren met ontwikkeling van COPD en jaarlijkse daling in FEV_1 in het Vlagtwedde/ Vlaardingen cohort. Aangezien *TIMP1* op het X-chromosoom ligt, hebben we mannen en vrouwen apart geanalyseerd; vrouwen dragen immers 2 kopieën van het gen en mannen 1. We vonden dat de *TIMP1* Ile158Val SNP vaker voorkwam bij vrouwen met COPD dan bij vrouwen zonder COPD. Ook was deze SNP geassocieerd met snellere daling in FEV_1 bij zowel mannen als vrouwen. De *TIMP1* SNP Phe124Phe was alleen bij mannen significant geassocieerd met snellere FEV_1 daling. De *MMP1* en *MMP12* SNPs waren niet geassocieerd met COPD of met longitudinale daling in FEV_1. We hebben vervolgens in een ander cohort van de algemene bevolking (het Doetinchem cohort) gekeken of de significante SNPs ook in een onafhankelijke algemene populatie geassocieerd waren met COPD en extra longfunctieverlies. Het bleek dat alleen de *TIMP1* Phe124Phe SNP geassocieerd was met longfunctieachteruitgang bij mannen. De andere twee associaties waren niet significant in het Doetinchem cohort, wat erop kan duiden dat het een foutpositieve bevinding was. Echter, het is ook mogelijk dat de associaties niet konden worden gerepliceerd door het feit dat de power om in het Doetinchem cohort significante associaties te vinden kleiner is dan in het Vlagtwedde/Vlaardingen cohort.

Niet alleen de afbraak van longweefsel kan van belang zijn in het ontstaan van COPD, maar ook een defect in het herstel van longweefsel. Een belangrijk gen hierbij is *transforming growth factor-β1* (*TGF-β1*) doordat het veel stoffen kan aantrekken die het beschadigde longweefsel herstellen. Een eiwit dat de functie van TGF-β1 kan remmen is decorine. Decorine expressie wordt vice versa gereguleerd door TGF-β1. Uit eerder onderzoek is gebleken dat TGF-β1 verhoogd en decorine verlaagd tot expressie komt in bronchiale biopten van patiënten met ernstig emfyseem. Het lijkt erop dat de balans tussen TGF-β1 en decorine verstoord is in COPD. Omdat het mogelijk is dat deze verstoorde balans veroorzaakt wordt door polymorfismen in *TGF-β1* en/of *decorine*, hebben we 3 SNPs in *TGF-β1* en 4 SNPs in *decorine* onderzocht in het

Vlagtwedde/Vlaardingen cohort (*hoofdstuk 7*). 1 SNP in *TGF-β1* (gelokaliseerd in de 3'UTR) kwam vaker voor bij mensen met COPD en een combinatie van twee SNPs (een zogenaamd haplotype) van twee andere SNPs in *TGF-β1* (C-509T en Leu10Pro) kwam significant minder vaak voor bij mensen met COPD. *Decorine* SNPs waren niet geassocieerd met ontwikkeling van COPD. We hebben ook geanalyseerd of SNPs in *TGF-β1* en *decorine* geassocieerd waren met extra daling in FEV_1 in de gehele populatie, maar we vonden geen associaties.

Effect van polymorfismen in antioxidant enzymen op longfunctieverlies en ontwikkeling van COPD en BHR

Sigarettenrook bevat veel zuurstofradicalen (oxidanten) die schadelijk zijn voor de luchtwegen en longen, en een ontstekingsreactie veroorzaken. Doordat de immuuncellen die hierdoor aangetrokken worden ook zuurstofradicalen produceren, ontstaat een situatie van chronische oxidatieve stress en ontsteking. De chronische ontsteking kan zorgen voor beschadiging van de longen, en bijdragen aan de ontwikkeling van COPD. In *hoofdstuk 8* en *9* hebben we onderzocht of genen die betrokken zijn in de reactie van lichaam tegen oxidatieve stress betrokken zijn bij de ontwikkeling van COPD.

Hoofdstuk 8 beschrijft de studie naar de zogeheten glutathion s-transferases (GSTs). Glutathion is een klein molecuul dat erg veel voorkomt in de longen en vrije radicalen kan neutraliseren. De GSTs zijn enzymen die ervoor zorgen dat glutathion ook kan binden aan grotere toxische stoffen in sigarettenrook. Er bestaan verschillende GST enzymen, waarvan wij *GSTP1, GSTM1* en *GSTT1* hebben onderzocht. Van de *GSTT1* en *GSTM1* genen bestaan twee bijzondere polymorfismen, namelijk deletie polymorfismen, wat betekent dat een deel van het gen uit het DNA is verdwenen, en het enzym dus ook niet gevormd kan worden. In *GSTP1* hebben we twee SNPs onderzocht. Het bleek dat zowel de deleties van *GSTM1* als *GSTT1* geassocieerd waren met verlies van longfunctie in de Vlagtwedde/Vlaardingen populatie. Omdat deze genen vooral betrokken zijn bij de neutralisatie van oxidanten in sigarettenrook, hebben we de genen verder onderzocht in rokers en niet-rokers apart. Het bleek dat rokers meer schade ondervonden van het verlies van deze enzymen dan niet-rokers, wat zich uitte in een groter longfunctieverlies bij rokers met de deletie dan rokers zonder de deletie polymorfismen. De SNPs in *GSTP1* waren niet geassocieerd met ontwikkeling van COPD en versnelde longfunctiedaling. Deze studie heeft laten zien dat een deficiëntie in enzymen betrokken bij neutralisatie van oxidanten in sigarettenrook een belangrijke bijdrage zou kunnen vormen bij de ontwikkeling van COPD bij rokers.

Er zijn meerdere enzymen betrokken bij het tegengaan van oxidatieve stress. De Superoxide Dismutases (SODs) neutraliseren zuurstofradicalen en verlagen daarmee de

oxidatieve stress. We hebben in *hoofdstuk 9* het effect van SNPs in *SOD-2* en *SOD-3* geanalyseerd op het verloop van longfunctie en de ontwikkeling van COPD. Aangezien oxidatieve stress ook een rol spelt in de ontwikkeling van bronchiale hyperreactiviteit (BHR) hebben we tevens geanalyseerd of deze SNPs ook gerelateerd waren aan de ontwikkeling van BHR in het Vlagtwedde/Vlaardingen cohort. De *SOD-2*C7693T SNP was significant geassocieerd met de ontwikkeling van zowel COPD als BHR in de gehele bevolking. Aangezien COPD en BHR sterk gecorreleerde fenotypes zijn hebben we groepen gemaakt van $COPD^{+}/BHR^{+}$, $COPD^{+}/BHR^{-}$, $COPD^{-}/BHR^{-}$ en $COPD^{-}/BHR^{+}$ personen. In deze groepen hebben we geanalyseerd of de SNPs geassocieerd waren met ofwel COPD of BHR. Deze analyses lieten zien dat de *SOD-2* C7693T SNP en de *SOD-2* Ala16Val SNP geassocieerd waren met BHR in personen zonder COPD. De *SOD-3* Arg213Gly SNP was geassocieerd met een langzamere daling in longfunctie in nooit rokers. Deze studie heeft aangetoond dat genetische variatie in *SOD* genen een belangrijke risicofactor zijn voor het ontstaan van COPD, en dat ze tevens belangrijk zijn in de ontwikkeling van BHR in de algemene bevolking.

VERANDERING IN SURFACTANT GENEN EN SURFACTANT LIPIDESAMENSTELLING IN ONTWIKKELING VAN COPD EN DALING VAN DE LONGFUNCTIE

Surfactant proteïnen (SFTP) spelen een rol in behoud van de structuur van de longblaasjes, in de eerste respons van het immuunsysteem op pathogenen en in regulatie van inflammatie in de longen. Een aantal polymorfismen in deze genen bleken in eerdere studies geassocieerd te zijn met hogere prevalentie van COPD. Wij hebben in *hoofdstuk 10* onderzocht of polymorfismen in *SFTPA1*, *SFTPA2*, *SFTPB* en *SFTPD* ook betrokken zijn bij versnelde achteruitgang in longfunctie. We vonden dat personen die heterozygoot zijn voor de *SFTPD* Met11Thr SNP vaker COPD hadden en ook een snellere achteruitgang hadden van de longfunctie (2,9 ml/jaar extra) vergeleken met personen met het wildtype genotype. Binnen personen met COPD daalden heterozygoten voor de *SFTPA1* Leu50Val SNP 11,6 ml/jaar minder hard dan personen met het wildtype genotype en deze hadden ook significant minder vaak COPD. Dit geeft aan dat SNPs in *SFTPA1* mogelijk slechts een functioneel effect kunnen hebben als er al een inflammatoir milieu is in de longen, zoals bij COPD patiënten het geval is.

Longsurfactant bestaat niet alleen uit eiwitten maar voor het grootste gedeelte uit vetten (lipiden). Om te onderzoeken of een verandering in samenstelling van deze lipiden ook samenhangt met de hoogte van de longfunctie en de ontwikkeling van COPD hebben we sputum van COPD patiënten en gezonde controles met elkaar vergeleken. De COPD patiënten hadden meer gedegradeerde lipiden in hun sputum dan de controles. De

hoeveelheid gedegradeerde lipiden was ook significant gecorreleerd met de hoogte van de longfunctie.

Uit deze studie is gebleken dat genetische variatie in surfactant genen bij zouden kunnen dragen aan de ontwikkeling van COPD, maar dat tevens de lipide samenstelling van belang is voor de hoogte van de longfunctie en ontwikkeling van COPD. Aangezien de groep met bekende lipidesamenstelling klein was, hebben we de relatie van genetische variatie en lipidesamenstelling niet kunnen onderzoeken. Dit is van belang voor vervolgonderzoek.

ONDERZOEK NAAR GENETISCHE FACTOREN VOOR ERNST VAN COPD IN EEN GENETISCH GE ̈ISOLEERDE POPULATIE

COPD wordt beschouwd als genetisch complexe ziekte, wat inhoudt dat er meerdere genen en omgevingsfactoren betrokken zijn bij het ontstaan van de ziekte. Het kan daarom nuttig zijn om ziektegenen op te sporen in een populatie waarin de genetische variatie en diversiteit in omgevingsfactoren beperkter is dan in de algemene populatie. Een genetisch ge ̈isoleerde populatie is hiervoor erg geschikt. De GRIP (Genetic Research in Isolated Populations) populatie is gelegen in het zuidwesten van Nederland bij de gemeente Rucphen. Deze populatie is gesticht rond het midden van de 18^{e} eeuw door ongeveer 150 personen en vrijwel ge ̈isoleerd gebleven tot een aantal decennia geleden. De populatie bestaat momenteel uit ongeveer 20,000 inwoners en van deze mensen is zeer gedetailleerd de familiegeschiedenis beschreven. Het is bekend dat veel inwoners verwant zijn via meerdere bloedlijnen en er bestaat een hoge mate van inteelt. Dit maakt dat de genetische diversiteit beperkt is en dus geschikt zou kunnen zijn voor genetisch onderzoek.

Voor het onderzoek beschreven in *hoofdstuk 11*, hebben we via de huisartsen van de gemeente Rucphen patiënten met een doktersdiagnose COPD gevraagd naar het onderzoekscentrum te komen om een longfunctietest te ondergaan, een vragenlijst te beantwoorden en bloed af te geven voor DNA isolatie. Dit resulteerde in een onderzoeksgroep van 111 personen met een matig tot ernstige vorm van COPD. Vervolgens hebben we SNPs in genen geselecteerd die in eerdere studies betrokken bleken te zijn bij de ontwikkeling van COPD, en gegenotypeerd in deze populatie. Het bleek dat een aantal van deze genen, namelijk *ADAM33*, *TGF-β1*, *SFTPA1* en *SFTPD*, geassocieerd waren met een lagere longfunctie bij patiënten met COPD. Omdat het door het isolement mogelijk is dat de onderzochte populatie genetisch gezien afwijkt van de algemene populatie, hebben we 351 personen geselecteerd uit het Vlagtwedde/Vlaardingen cohort met een gelijke ernst van COPD. Vervolgens hebben we geanalyseerd of de significant geassocieerde genen ook in dit replicatiecohort geassocieerd waren met longfunctie. Dit bleek zo te zijn voor *ADAM33* en *TGF-β1*, maar

niet voor *SFTPA1* en *SFTPD*. Deze laatste twee associaties zouden dus toevalsbevindingen kunnen zijn. Deze studie laat ons zien dat genetisch geïsoleerde populatie van GRIP gebruikt kan worden om in een relatief kleine onderzoekspopulatie genen op te sporen die geassocieerd zijn met de ernst van COPD.

BELANGRIJKSTE CONCLUSIES VAN DIT PROEFSCHRIFT

Onze resultaten hebben laten zien dat genetische factoren geassocieerd zijn met de ontwikkeling van COPD en de ernst van deze ziekte. Aangezien er nog steeds geen goede medicatie is voor COPD patiënten zijn studies zoals door ons uitgevoerd belangrijk om het mechanisme achter het ontstaan van COPD beter in kaart te brengen en wellicht om nieuwe therapieën te ontwikkelen.

De belangrijkste conclusies van dit poefschrift zijn:

1. Het *ADAM33* gen speelt een rol in de ontwikkeling van COPD en versnelde longfunctiedaling in de algemene populatie, en is geassocieerd met de ernst van COPD
2. Genetische variatie in de proteases *MMP1*, *MMP2*, *MMP9* en *MMP12* draagt niet bij tot een verminderde longfunctie, genetische variatie in de remmer van deze proteases, *TIMP1,* echter wel
3. *TGF-β1* is geassocieerd met ontwikkeling van COPD en ernst van COPD, maar decorine niet
4. Deleties van de *GSTM1* en *GSTT1* enzymen, die de rookgerelateerde oxidatieve stress verminderen, zijn geassocieerd met versnelde longfunctiedaling in rokers
5. Polymorfismen in de genen van de *Superoxide Dismutase* enzymen zijn betrokken bij de aanwezigheid van bronchiale hyperreactiviteit en achteruitgang van longfunctie in de algemene bevolking
6. Zowel de samenstelling van lipiden in long surfactant als genetische varianten in surfactant genen zijn risicofactoren voor de ontwikkeling van COPD en het niveau en verloop van longfunctie
7. Genetisch geïsoleerde populaties kunnen in genetisch onderzoek naar COPD gebruikt worden voor identificatie van genen betrokken bij de ernst van COPD

Dankwoord

Acknowledgements

DANKWOORD

Promoveren doe je niet alleen. Ik wil graag iedereen bedanken die op welke manier dan ook heeft bijgedragen aan de totstandkoming van dit proefschrift. Een aantal mensen wil ik hierbij in het bijzonder noemen.

Dit proefschrift zou er niet zijn geweest zonder de grote inzet en betrokkenheid van de deelnemers aan het Vlagtwedde/Vlaardingen onderzoek en alle medewerkers die hebben bijgedragen aan de metingen, dataverzameling en databeheer. Ik vind het erg bijzonder dat we 18 jaar na beëindiging van de oorspronkelijke studie nog steeds zo goed gebruik kunnen maken van de data. Hartelijk dank!

Mijn grote dank gaat uit naar mijn begeleiders in mijn onderzoek, mijn copromotor Dr. Marike Boezen en mijn promotor Prof. Dr. Dirkje Postma.
Beste Marike, 4 jaar begeleiding is moeilijk samen te vatten in een paar zinnen. Natuurlijk heb ik veel van je geleerd over de epidemiologische aanpak van onderzoek en je kritische blik op manuscripten en presentaties. Ik waardeer je nuchtere en relativerende kijk op onderzoek. Je vertrouwen in mij en het verloop van het onderzoek heb ik erg op prijs gesteld en heeft mij in staat gesteld om goed zelfstandig onderzoek te doen, wat ik goed zal kunnen gebruiken in mijn verdere loopbaan. Veel dank!
Beste Dirkje, de afgelopen jaren heb je me duidelijk gemaakt dat het in onderzoek belangrijk is om het grote geheel te zien en buiten je eigen specifieke onderzoeksgebied te blijven kijken. Het is eenvoudig te verzanden in analyseren van grote databestanden en een associatie heeft weinig betekenis als je die niet probeert te betrekken in de pathofysiologie van een ziekte. Van deze benadering heb ik veel geleerd. Ook heb ik erg je enthousiasme over de resultaten gewaardeerd. Heel erg bedankt!

I would like to thank the members of the reading committee, Prof. dr. Gerard Koëter, Prof. dr. Cisca Wijmenga, Prof. dr. Peter Paré, for their willingness and time to review the manuscript.

Alle collega's van de afdeling Epidemiologie wil ik hartelijk bedanken voor de prettige werksfeer en gezellige borrels en uitjes.
Graag wil ik in het bijzonder de epi-ladies en heer vanaf het begin van mijn onderzoek bedanken: Judith, Marjan, Marike, Désirée en Jan. Jullie zijn doorgewinterde epidemiologen, en ik heb veel van jullie opgestoken over onderzoeksopzet en statistiek. Bedankt voor al jullie hulp, interesse, ins en outs over het Vlagtwedde/Vlaardingen cohort en natuurlijk de gezelligheid! Judith, jou wil ik bedanken voor de prettige manier waarop je me hebt geholpen met de LME's, de SPSS trucjes, en de meerwaarde van het gebruik de syntax. Het was fijn steeds bij je te kunnen komen voor kleine en grote vragen, en het was erg leuk een kamer met je te delen.
Mateusz, it was very nice having you as a roommate, I really enjoyed our discussions about our projects and the more general ones. Thank you very much for that. I wish you all the best with your project, and hopefully we can work together again sometimes.
Hiltje en Mireille, bedankt voor de vele leuke gesprekjes over het wel en wee van werk, vette levers en vette kindertjes, promoveren, katten, huizen, en natuurlijk de laatste

nieuwtjes. Ik vind het heel fijn dat jullie mijn paranimfen zijn, en ik wens jullie heel veel succes met jullie promotieonderzoek.
Greetje, bedankt voor de gezellige gang-gesprekjes. Ilja, hartelijk dank voor je hulp met de haplotype-analyses in Arlequin.

Collega's van het 'oude genotyperingslab': Marcel, Elvira, Matthieu, Gerrit, Elinda en Niels, hartelijk dank voor de erg prettige samenwerking op het lab. In het bijzonder wil ik Marcel bedanken voor het kennismaken met de genotypering en je immer aanwezige bereidwilligheid om mee te denken. Elvira, bedankt voor je praktische hulp op het lab.

Wim van der Bij en Gerard Koëter wil ik hartelijk danken voor het initiëren van de genetische studie naar de longtransplantatiepatiënten en de betrokkenheid en interesse bij het verdere verloop van de studie. Voor deze studie zijn oude longfunctiegegevens opgevraagd en ingevoerd door Willy, bedankt voor je werk. De immunohistochemie voor deze studie is uitgevoerd op het lab van de Pathologie onder begeleiding van Wim Timens. Beste Bea, hartelijk dank voor het snijden van de vele coupes en alle hulp bij de kleuringen. Beste Wim, bedankt voor de uitleg en het overleg over de immunohistochemie.

Voor de *ADAM33* studie in de GLUCOLD populatie hebben we er vele analysemiddagen opzitten, maar saai waren ze nooit. Ik wil Margot Gosman bedanken voor de fijne samenwerking! Alle co-auteurs, dank voor de kritische blik op het manuscript.

I would like to thank Begona Barroso and Rainer Bisschoff for their collaboration, the explanation of the results of the lipidomics and critically reading of the chapter on surfactant.

In 2005 hebben we longfunctiegegevens verzameld van COPD patiënten in de genetisch geïsoleerde bevolking van Rucphen in het kader van de GRIP studie. Ik wil de deelnemers hartelijk danken voor hun medewerking, de longfunctie assistenten voor de gezellige tijd tijdens het veldwerk, Leon Testers voor de goede coördinatie ter plaatse en met het verzenden van de DNA'tjes na afloop. Yurii Aulchenko, thank you very much for spending time in explaining the complex statistics required for analyzing pedigree data and for your help analyzing and interpreting the data. Cock van Duijn, hartelijk dank voor de samenwerking en je kritische blik op het manuscript. Alle co-auteurs, bedankt voor de samenwerking.

Het Nederlands Astmafonds wil ik hartelijk danken voor het financieren van het project. Het heeft het mogelijk gemaakt weer een stapje verder te komen in het begrijpen waarom sommige mensen COPD krijgen en anderen niet.

Onderzoek moet je niet alleen willen doen, en het bespreken van resultaten draagt zeker bij aan beter begrijpen van de materie. Ik wil de deelnemers van maandelijkse GRIAC Genetica vergadering bedanken voor de discussies, presentaties en samenwerking.

Niet alleen wordt er veel genetica besproken binnen het GRIAC, maar minstens zoveel heb ik geleerd van de GRIAC dinsdagmiddag besprekingen. GRIAC'ers, heel erg bedankt voor de presentaties, commentaren, suggesties en de leuke congressen.

Ik wil Jeroen Advocaat ontzettend bedanken voor de mooie omslag en de prettige, vlotte samenwerking!

Mijn familie en vrienden wil ik bedanken voor de interesse en afleiding tijdens mijn onderzoek, zoals het real life experiment van het combineren van Eising en Van Diemen genen, wat heeft geresulteerd in het aller-leukste nichtje dat er maar is.
Bas en Thierry, vrienden vanuit Leiden, maar nu helaas nog verder weg wonend, ik vind het erg bijzonder dat onze vriendschap in stand blijft, bedankt!
Lieve mam, promoveren is weer een behaald station; heel erg bedankt voor de ruimte die je me altijd hebt gegeven om zelf te kijken waar ik heen wilde.

En tenslotte Wouter: niet teveel poeha hoor, maar toch bedankt voor zo'n fijn thuis!

Cleo